Gray-box Modeling for Stable and Efficient Operation of Steel Making Process(Dissertation_全文)

AUTHOR(S):

Ahmad, Iftikhar

Gray-box Modeling for Stable and Efficient Operation of Steel Making Process

Iftikhar AHMAD

2014

Contents

Chapter 1

Introduction

1.1 Steel Making

Steel is the most important material for construction and consumer products, and its manufacturing industry is one of the largest energy consumer in the world. Steel is an alloy made by adding carbon and several alloys to molten iron followed by chemical treatments in a highly heated environment. Comparatively low cost of production than the other alloys, i.e., aluminum alloys etc., the abundant availability of main raw material (iron ore) and diverse range of mechanical properties lead to its large scale of production. The alloying elements and the heat treatment determine the properties of the steel products, and developments in these two aspects have been the main drive of evolution of the steel making industry.

1.1.1 Evolution of Steel Making Industry

The evolution of the modern day steel making industry trace back to the use of iron during the 2500-2000 BC, mainly obtained from meteorites [1]. In 1300 BC, the use of fuels such as charcoal begun to heat the iron for conversion into wrought iron. However the achievable temperature was far

below the melting point of iron and therefore hammering was used after heating for its conversion to wrought iron. Since that time till date, search for better heating technique has been one of the main focus of the users. This drove the evolution of heating system with passage of time and at the 1300 AD, the induction of Stuckoven process laid the foundation of the modern day blast furnace used in steel making. Around 1850 AD, the puddle process which had been used for iron making for a century, was converted to steel making unit. However, the product was still semi-fluid and needed forging and therefore the technique could not get enough success. In 1860, the first Bessemer process was launched. It could oxidize carbon and other impurities and reduced the refining time and costs. Still, the remnants of sulfur, phosphorus and large quantities of nitrogen dissolved in the steel, caused low product quality. The Thomas process, an adapted version of the Bessemer process, and open hearth furnace (OHF) were introduced in 1878. Those lead to the production of low phosphorus steel from high phosphorus pig iron and it was possible to get sufficiently high temperature to melt steel. In 1952, the basic oxygen furnace (BOF) was introduced with far better energy efficiency and faster refining than the OHF. It realized the idea of using pure oxygen mentioned by Bessemer in 1856. Later on, the induction of the electric arc furnace (EAF) provided a completely different route to steel making, by achieving high temperature and accepting all types of input iron, even 100% scrap.

The evolved modern day steel making industry is equipped with highly efficient heating and other treatment units. Process flow diagram of the steel making process is shown in **Figure 1.1**. This dissertation focus on modeling the molten steel temperature at steel making from the converter to the continuous casting process. Modeled sections of the steel making process in this dissertation are shown in **Figure 1.2**. The molten steel in the converter is discharged to a ladle, and in order to adjust the composition, ferro alloys are added to the molten steel. Then, the molten steel is transported to

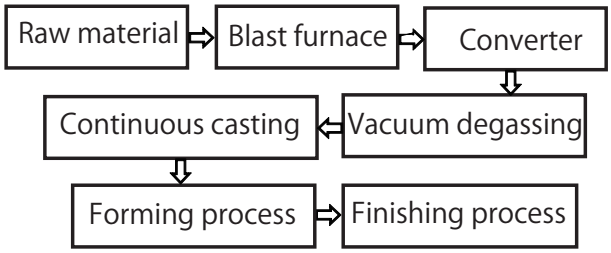


Figure 1.1: Process flow diagram of the steel making process

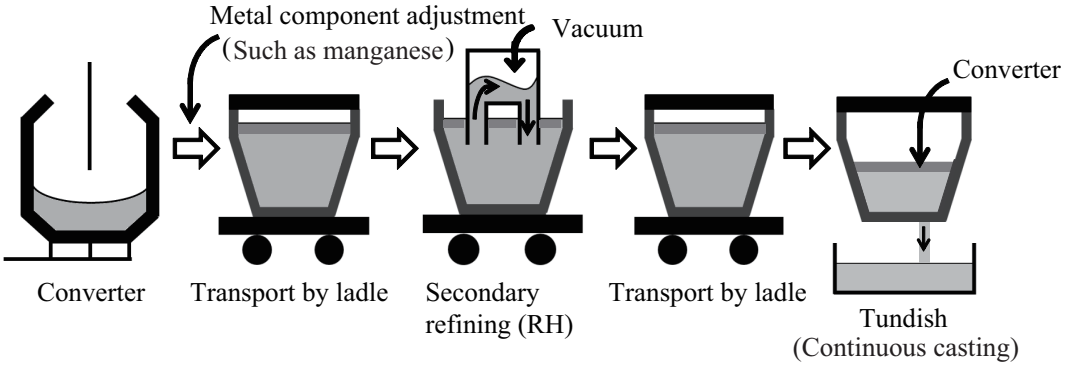


Figure 1.2: Modeled sections of the steel making process

the secondary refining process called the Ruhrstahl-Heraeus (RH) process. After the secondary refining, the molten steel is transported to the continuous casting process, and is discharged to the tundish which is a vessel used for delivering molten steel from a ladle to a mold in the continuous casting process.

1.1.2 Modeling through Statistical and Computational Techniques

The modern qualitative testing facilities and new trends in computational modeling make a way to strive for achieving a more efficient steel making process. The use of statistical and computational techniques to model the complex steel making processes and make further insight to understand, de-

sign, and control of these processes enhances competition within the steel making industries in the global market. Phenomenon such as transient behavior during steady casting, interfacial slag layer heat, mass and momentum balances, and non-equilibrium crystallization behavior, and slag structure and friction are modeled both through the simulation techniques as well as the real plant data [2].

Control of the molten steel temperature in the tundish (TD temp) is one of the key factors to realizing stable operation in steel making. If TD temp is too high, breakouts may occur and cause tremendous increase in maintenance cost and productivity loss. When the temperature is too low, clogging in the tundish nozzle occurs, which causes disruptions in the casting process. However, no effective manipulated variable is available after the secondary refining process to control TD temp. To realize the target TD temp, therefore, it is necessary to adjust the molten steel temperature in the secondary refining process (Ruhrstahl-Heraeus degassing process), RH temp. To control TD temp by manipulating RH temp, a model relating TD temp and RH temp needs to be constructed. In the past, various models such as first-principle models [3]−[8], statistical models [9], and gray-box models [10] have been proposed to predict the molten steel temperature.

The gray-box model, which integrates a first-principle model and a statistical model, has attracted researchers' attention by its capability: known linear/nonlinear phenomena can be embedded in the first-principle model, and an unknown relationship among variables can be embedded in the statistical model by extracting such a relationship from the data. In general, gray-box models are more accurate than simplified first-principle models, less complicated than computational fluid dynamics (CFD) models, and more easily interpreted than statistical models. Although a gray-box model aims to improve the prediction performance by combining a first-principle model and a statistical model, the accuracy of the first-principle model is still important. In general, first-principle models have various parameters which need

to be determined by using data. Even when some parameters depend on the operating conditions, they are kept constant if it is difficult to identify the relationship between the parameters and the operating conditions. In such a case, large prediction errors might be caused. Such errors in identifying a relationship between the parameters and operation condition is always a challenge to modeling a complex process. The reasons of the prediction errors is uncertainties in equipment characteristics, operating conditions, and raw materials. In this research, new gray-box modeling technique is proposed to minimize prediction errors caused by such uncertainties.

1.2 Gray-box Models

In this section, gray-box models are categorized into two types, parallel gray-box models [11] and serial gray-box models [12]. A generalized framework of the two types of gray-box model is explained.

1.2.1 Parallel Gray-box Model

A typical gray-box model is constructed by combining a first-principle model and a statistical model sequentially; the statistical model is built so as to compensate the errors of the first-principle model. This type of gray-box model, called the parallel gray-box model, is developed through the following steps.

i. Build a first-principle model f_{fp} to predict an output variable y from input variables X_{fp}.

$$\hat{y}_{\mathrm{fp}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \boldsymbol{\theta}) \tag{1.1}$$

where $\hat{y}_{\mathrm{fp}}$ is the prediction of y and $\boldsymbol{\theta}$ is a parameter vector.

ii. Estimate the parameters $\boldsymbol{\theta}$ by solving the optimization problem of minimizing the sum of squared errors.

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{n=1}^{N_{\mathrm{d}}} e_{\mathrm{fp},n}^2 \tag{1.2}$$

$$e_{\mathrm{fp},n} = y_n - f_{\mathrm{fp}}(X_{\mathrm{fp},n}, \boldsymbol{\theta}) \tag{1.3}$$

$$\boldsymbol{\theta}_{\mathrm{L}} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\mathrm{U}} \tag{1.4}$$

where $X_{\mathrm{fp},n}$ and y_n are the nth sample of input and output variables, respectively. N_{d} is the number of samples used for developing the model. $\boldsymbol{\theta}_{\mathrm{L}}$ and $\boldsymbol{\theta}_{\mathrm{U}}$ are lower and upper bound vectors of parameters which are determined in advance.

iii. Build a statistical model f_{pa} to predict the output errors e_{fp} from input variables X by solving the following optimization problem.

$$\tilde{\boldsymbol{\varphi}}_{\mathrm{pa}} = \underset{\boldsymbol{\varphi}_{\mathrm{pa}}}{\operatorname{argmin}} \sum_{n=1}^{N_{\mathrm{d}}} (e_{\mathrm{fp},n} - f_{\mathrm{pa}}(X_n, \boldsymbol{\varphi}_{\mathrm{pa}}))^2 \tag{1.5}$$

$$\hat{e}_{\mathrm{fp},n} = f_{\mathrm{pa}}(X_n, \boldsymbol{\varphi}_{\mathrm{pa}}) \tag{1.6}$$

where $\boldsymbol{\varphi}_{\mathrm{pa}}$ is parameters in the statistical model. In general, X_{fp} is a subset of X.

iv. Build a gray-box model by combining the first-principle model and the statistical model.

$$\hat{y}_{\mathrm{pa}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \tilde{\boldsymbol{\theta}}) + f_{\mathrm{pa}}(X, \tilde{\boldsymbol{\varphi}}_{\mathrm{pa}}) \tag{1.7}$$

where $\hat{y}_{\mathrm{pa}}$ is the prediction of y by using the parallel gray-box model.

1.2.2 Serial Gray-box Model

In the parallel gray-box model, the optimal parameters $\tilde{\boldsymbol{\theta}}$ are constant, although some parameters depend on the operating conditions. This simplification may deteriorate the prediction performance of the first-principle model. Thus, another type of gray-box model, called the serial gray-box model, is used to estimate the parameters as functions of input variables [12]. The serial gray-box model is constructed by the following steps.

i. A first-principle model $\hat{y}_{\mathrm{fp}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \boldsymbol{\theta})$ is developed.

ii. The parameters $\boldsymbol{\theta}$ in the first-principle model are optimized by Eq. (1.2).

iii. Certain parameter/parameters θ is/are optimized for each modeling sample.

$$\tilde{\theta}_n = \underset{\theta_{i,n}}{\operatorname{argmin}} (y_n - f_{\mathrm{fp}}(X_{\mathrm{fp},n}, \tilde{\boldsymbol{\theta}}^c, \theta_n))^2 \ (n = 1, 2, \cdots, N_{\mathrm{d}}) \tag{1.8}$$

$$\theta_{\mathrm{L}} \leq \theta_n \leq \theta_{\mathrm{U}} \tag{1.9}$$

where $\tilde{\theta}_n$ is the optimal value of the parameter θ for sample n and $\tilde{\boldsymbol{\theta}}^c$ is the constant vector consisting of the parameters except θ. N_p is the number of parameters in the first-principle model. θ_L and θ_U are lower and upper bounds of the parameter which are determined in advance.

iv. A statistical model f_se is constructed to estimate the parameter θ from the input variables X by solving the following optimization problem.

$$\tilde{\boldsymbol{\varphi}}_\mathrm{se} = \underset{\boldsymbol{\varphi}_\mathrm{se}}{\mathrm{argmin}} \sum_{n=1}^{N_\mathrm{d}} (\tilde{\theta}_n - f_\mathrm{se}(X_n, \boldsymbol{\varphi}_\mathrm{se}))^2 \tag{1.10}$$

where $\tilde{\boldsymbol{\varphi}}_\mathrm{se}$ represents the optimal values of parameters.

v. By combining the first-principle model and the statistical model, the serial gray-box model is derived:

$$\hat{y}_\mathrm{se} = f_\mathrm{fp}(X_\mathrm{fp}, \tilde{\boldsymbol{\theta}}^c, \hat{\tilde{\theta}}) \tag{1.11}$$

$$\hat{\tilde{\theta}} = f_\mathrm{se}(X, \tilde{\boldsymbol{\varphi}}_\mathrm{se}) \tag{1.12}$$

1.3 Statistical Models

1.3.1 Partial Least Squares (PLS)

PLS is a linear regression method that can cope with the collinearity problem; thus it has been used as a modeling tool in various industries where process variables are highly correlated.

In PLS with one output variable, input data $\boldsymbol{X} \in \Re^{N \times M}$ and output data $\boldsymbol{y} \in \Re^{N}$ are decomposed as follows:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E} \tag{1.13}$$

$$\boldsymbol{y} = \boldsymbol{T}\boldsymbol{b} + \boldsymbol{f} \tag{1.14}$$

where $\boldsymbol{T} \in \Re^{N \times R}$ is a latent variable matrix whose columns are latent variables $\boldsymbol{t}_r \in \Re^N$ $(r = 1, 2, \cdots, R)$, $\boldsymbol{P} \in \Re^{M \times R}$ is a loading matrix of $\boldsymbol{X}$ and its columns are loading vectors $\boldsymbol{p}_r$, $\boldsymbol{b} = [b_1, b_2, \cdots, b_R]^T$ is a loading vector of $\boldsymbol{y}$, and $\boldsymbol{E}$ and $\boldsymbol{f}$ are errors. N, M, and R denote the number of samples, that of input variables, and that of adopted latent variables, respectively.

The nonlinear iterative partial least squares (NIPALS) algorithm can be used to construct the PLS model. Suppose that the first to $(r-1)$th latent variables $\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_{r-1}$, the loading vectors $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_{r-1}$ and $b_1, b_2, \cdots$, b_{r-1} are given. The rth residual input and output can be written as follows:

$$\boldsymbol{X}_r = \boldsymbol{X}_{r-1} - \boldsymbol{t}_{r-1}\boldsymbol{p}_{r-1}^T \tag{1.15}$$

$$\boldsymbol{y}_r = \boldsymbol{y}_{r-1} - b_{r-1}\boldsymbol{t}_{r-1} \tag{1.16}$$

where $\boldsymbol{X}_1 = \boldsymbol{X}$ and $\boldsymbol{y}_1 = \boldsymbol{y}$. The latent variable $\boldsymbol{t}_r$ is a linear combination of the columns of $\boldsymbol{X}_r$, that is, $\boldsymbol{t}_r = \boldsymbol{X}_r\boldsymbol{w}_r$ where $\boldsymbol{w}_r \in \Re^M$ is the rth weighting vector. PLS aims to maximize the covariance between $\boldsymbol{y}_r$ and $\boldsymbol{t}_r$ under the constraint $\|\boldsymbol{w}_r\| = 1$. By using the Lagrange multipliers method, $\boldsymbol{w}_r$ is derived as

$$\boldsymbol{w}_r = \frac{\boldsymbol{X}_r^T\boldsymbol{y}_r}{\|\boldsymbol{X}_r^T\boldsymbol{y}_r\|} \tag{1.17}$$

The rth loading $\boldsymbol{p}_r$ and b_r are as follows:

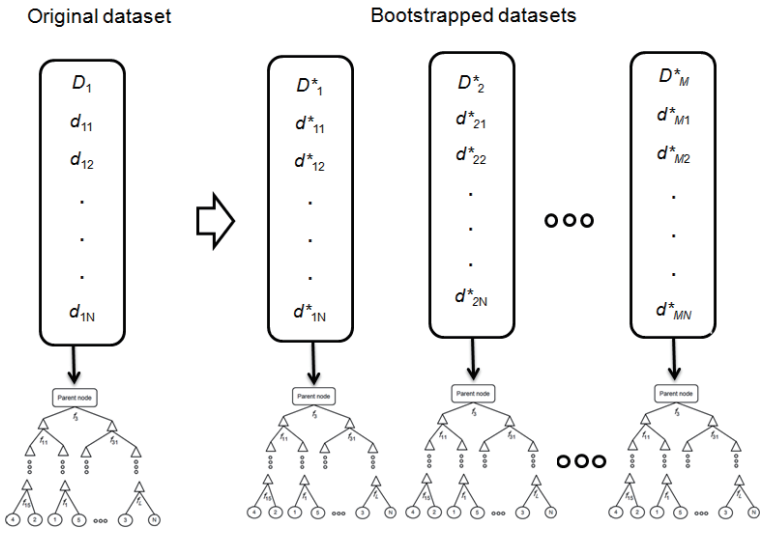
$$\boldsymbol{p}_r = \frac{\boldsymbol{X}_r^T \boldsymbol{t}_r}{\boldsymbol{t}_r^T \boldsymbol{t}_r} \tag{1.18}$$

$$b_r = \frac{\boldsymbol{y}_r^T \boldsymbol{t}_r}{\boldsymbol{t}_r^T \boldsymbol{t}_r} \tag{1.19}$$

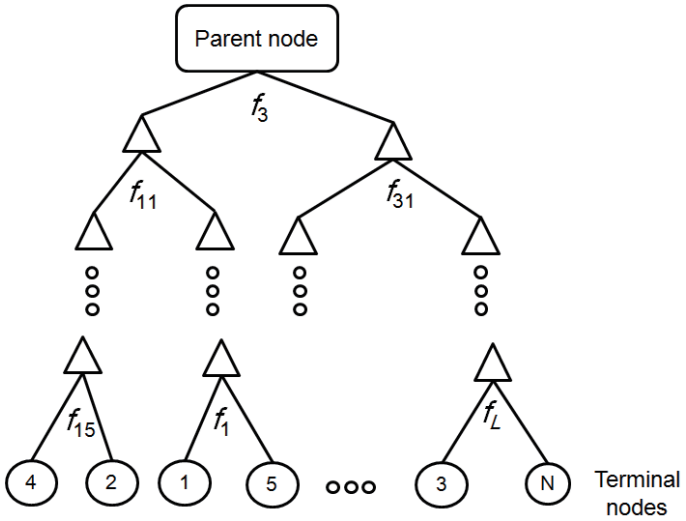
Finally, the above procedure is repeated until the number of adopted latent variables R is achieved; R can be determined by some cross validation technique [13]−[14]. In cross validation, training data is divided into several subgroups. PLS models are constructed iteratively. At each iteration, one subgroup is used for model validation and the other subgroups are used for model construction. On the basis of the sum of prediction errors, the optimal number of latent variables adopted in the PLS model is determined.

1.3.2 Random Forests (RF)

Random forests (RF) is an ensemble classifier that consists of many decision trees [15]. RF combines Brahman's bagging idea and the random selection of split features [16]−[17]. Given a training set D of size N, bagging generates J new training sets $D_i^*(i = 1, 2, \cdots, L)$, whose size is N, by random sampling from D with replacement. RF creates multiple decision trees; each tree is trained by using the bootstrapped samples as shown in **Figure 1.3(a)**. In construction of the tree, Figure 1.3(b), at each node of the tree, feature variables, i.e., split features, are randomly selected and splitting is performed using these features one by one to find the best split. For each terminal node, the average of the output variable is assigned. When a new sample is given, it is classified to a terminal node for each tree, and the value of output variable is obtained. The estimated value of the output variable for the new sample is given as the average of these values.



(a) Development of bootstrap samples



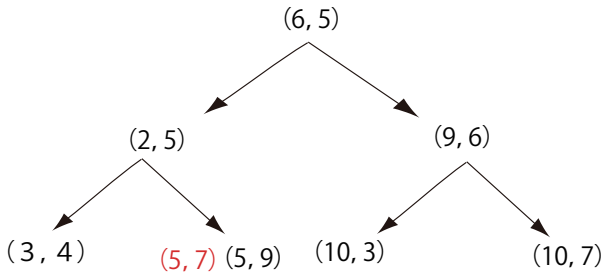
(b) Construction of decision tree

Figure 1.3: Construction of random forests

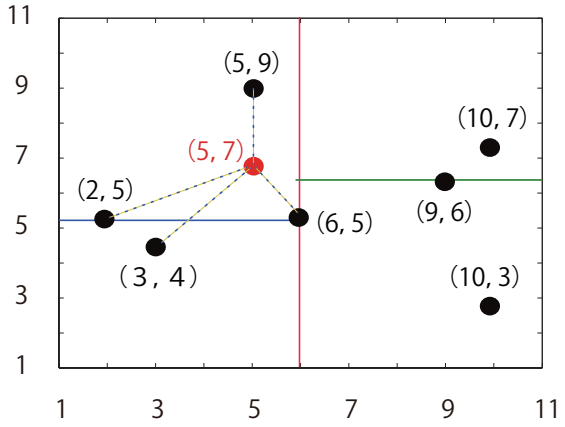
1.4 K-D Tree

The k-d tree is a data structure used for organizing data in a d dimensional space, and nearest neighbor search [18]. Given a dataset of n samples with dimensions (feature variables) f_i ($i = 1, 2, \cdots, L$), initially a dimension f_i is selected to split the dataset into two subsets, i.e., left child subsets and right child subsets. The selection of f_i is based on its largest spread compared to rest of dimensions in L. The median of f_i is used as splitting threshold; the samples who's f_i is smaller than the threshold go to the left child subset while those having larger f_i than the threshold go to the right child subset. Then in the next step, another dimension is selected to split the two child subsets into further two child subsets, each. This splitting continuous recursively until a child subset with just one sample is left. The structure of the constructed tree is then used to find k nearest neighbors of a given query q. Search for the nearest neighbors is started at the root node which proceeds recursively through all the nodes. At each node, the same dimension f_i which was used in construction of the k-d tree is used to decide direction of search for the nearest neighbors [19].

Let's suppose a dataset containing seven samples (3,4), (2,5), (5,9), (6,5), (9,6), (10,7), and (10,3) with two dimensions, i.e., X and Y. The construction of k-d tree for this dataset is shown in **Figure 1.4**. The X dimension being of the largest spread is selected for spiting the dataset into two child subsets. The X dimension value, 6, of sample (6, 5) is used as a median, although it is not the exact median but closest to the actual median of the X dimension. In the next split Y dimension is selected for splitting the child subsets into further child subsets, where the respective median at each child subset is used as splitting thresholds. Then to find the nearest neighbor of a query sample (5, 7), the X and Y dimensions are used sequentially to reach the nearest sample (5, 9) according to tree structure in Figure 1.4 (a). To make sure that (5, 9) is the nearest neighbor sample, Euclidean distance between the query and all the samples on the left side of the k-d tree are calculated as shown in



(a) The resulting *k-d* tree structure



(b) The *k-d* tree decomposition for the given dataset

Figure 1.4: The *k-d* tree construction

Figure 1.4 (b). In this example, (5, 9) is the nearest sample according to the Euclidean distance as well. However, in case the other samples have smaller distance than (5, 9), then the earlier result is discarded and the new sample is considered as the nearest neighbor sample.

1.5 Bootstrap Filter

Bootstrap filter is a filtering technique to represent and recursively generate an approximation to the state probability density function (PDF) as a set of random samples [20]. The bootstrap filtering method can handle the issue of non-linearity or non Gaussian noise. Use of a sufficiently large number of random samples effectively provide the required PDF. Let's suppose a state space model describing the process dynamics consists of a state transition model and an observation model. The transition of target variables y from y_{i-1} to y_i is described by

$$y_i = f_i(y_{i-1}, X_i) + \varepsilon_i \tag{1.20}$$

where f_i is a deterministic transition function, i represents sequential phases in the process, X_i is the input variables and ε_i is a system noise. Assuming that ε_i is Gaussian with mean 0 and variance $\sigma^2_{\varepsilon,i}$,

$$P(y_i|y_{i-1}, X_i) = N(f_i(y_{i-1}, X_i), \sigma^2_{\varepsilon,i}) \tag{1.21}$$

where $N(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2. The observation $\tilde{y}_i$ of y_i is modeled with an observation errors ξ_i, which is also Gaussian with mean 0 and variance $\sigma^2_{\xi,i}$.

$$\tilde{y}_i = y_i + \xi_i \tag{1.22}$$

$$P(\tilde{y}_i|y_i) = N(y_i, \sigma^2_{\xi,i}) \tag{1.23}$$

To develop a bootstrap algorithm, a certain number of samples N is initialized and for each of these samples, the likelihood of $y_i^{(n)}$ to the measured y_i is calculated:

$$W_i^{(n)} = p((y_i^{(n)} - \tilde{y}_i), \sigma_i^2), \tag{1.24}$$

where $p(x)$ is a probability density function of the probability distribution. Then $\{y_i^{'(n)}\}_{n=1}^N$ is resampled from $\{y_i^{(n)}\}_{n=1}^N$ with replacement in such a way that the probability of the frequency of $y_i^{(n)}$ is proportional to $W_i^{(n)}$. For all the n resampled samples, $y_{i+1}(n)$ is predicted by Eq. (1.20) and proceeded to phase $i+1$. The process continue until the probability distribution of the output variable of the last phase of the process is predicted. For a three phase process, i.e., $i = 1, 2, 3$, the transformation of prior distribution to posterior distribution for a target variable is shown in **Figure 1.5**. Given an adequate number of samples, the bootstrap filter outperforms the Kalman filters [20].

1.6 Objective and Thesis Structure

The objective of this dissertation is to propose a systematic method to develop a model for prediction and control of molten steel temperature in continuous casting process. The work consists of the following chapters: chapter 2 describes a model developed for prediction of molten steel temperature in tundish. A first-principle model was developed on the basis of computational fluid dynamics (CFD) simulations to simplify the model, then to improve the estimation accuracy statistical models were developed to estimate the estimation errors of the first-principle model through partial least squares (PLS) and random forests (RF). As a result of comparing the models, the

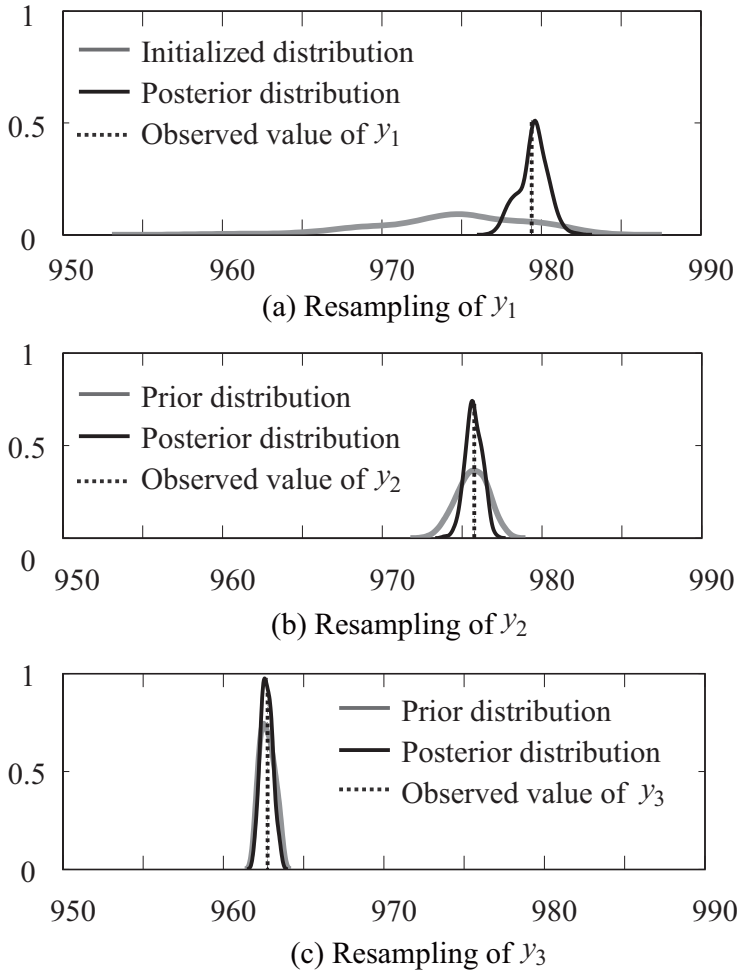


Figure 1.5: Prior distributions, posterior distributions and observed values of target variable y

RF-based gray-box model achieved the best estimation performance. The parallel gray-box model can extract information from data that is not used in the first-principle model and also it can overcome the limitations imposed by the structure of the first-principle model. However, it does not investigate the outdated parameters of the first-principle model and simply compensate the errors, by the statistical models. The outdated parameters should be searched and updated according to the changing operation condition. To overcome this flaw, another gray-box model is proposed in the next chapter.

Chapter 3 describes another model proposed to improve the prediction efficiency of the gray-box model developed in chapter 2, and make it able to control molten steel temperature in tundish. The overall heat transfer coefficient of a ladle, which is a parameter in the first-principle model, is optimized for each past batch separately, then the relationship between the optimal values of the parameter and measurements of process variables is modeled through a statistical modeling method, RF. The statistical model updates the parameter according to the operating condition. Such a model, including a statistical model inside a first principle model, is termed as a serial gray-box model. The serial gray-box modeling lead to better understanding of the process by recursively searching the most effective parameters, and optimize them to minimize the effect of changing operating conditions on the prediction performance of the first-principle model. Although, the serial gray-box modeling approach update the selected parameters and minimize the prediction errors better than the parallel gray-box modeling, there is still a significant amount of errors left. To overcome this remaining errors, another model called the combined gray-box modeling is proposed which outperforms all the other models, i.e., the first-principle model, the statistical model, the parallel gray-box model and the serial gray-box model. In addition, the proposed gray-box model is used to determine the molten steel temperature in the Ruhrstahl-Heraeus degassing process (RH degasser) at the end of its operation from the target TD temp.

In chapter 4, the issue of process uncertainties caused by equipment characteristics, operating conditions, and raw materials are taken into account. Due to lack of process information, realizing such mathematical expressions for process uncertainties is always challenging and has not been much focused. In this chapter a model is proposed by embedding the combined gray-box model with a bootstrap filter. The combined gray-box model is used as a deterministic function for the bootstrap filter. General shaped probability distributions of the molten steel temperature can be computed with the particle approximation technique. Furthermore, to minimize the risk caused by such uncertainties and realize the optimal operation, the reliability of prediction is evaluated on the basis of its probability distribution. Chapter 5, concludes the whole work proposed in this dissertation.

Bibliography

[1] J. d. Beer, E. Worrell, K. Blok: Future technologies for energy-efficient iron and steel making, *Annu. Rev. Energy Environ*, 23 (1998), 123-205.

[2] B. G. Thomas: Computational modeling of flow, heat transfer, and deformation in the continuous casting of steel, *1st Baosteel Annual Academic Conference*, May 27-28, Shanghai, China, (2004).

[3] A. Zoryk, P. M. Reid: On-line liquid steel temperature control, *Iron and Steelmaker*, 20 (1993), 21-27.

[4] A. G. Belkovskii, Y. L. Kats: Mathematical model of the cooling of steel in a small ladle, *Metallurgist*, 53 (2009), 261-273.

[5] P. R. Austin, J. M. Camplin, J. Herbertson, I. J. Taggart: Mathematical modeling of thermal stratification and drainage of steel ladles, *ISIJ International*, 32 (2) (1992), 196-202.

[6] J. L. Xia, T. Ahokainen: Transient flow and heat transfer in a steelmaking ladle during the holding period, *Metallurgical and Materials Transactions B*, 32 (2001), 733-741.

[7] J. R. S. Zabadal, M. T. M. B. Vilhena, S. Q. B. Leit: Heat transfer process simulation by finite differences for online control of ladle furnaces, *Ironmaking and Steelmaking*, 31 (2004), 227-232.

[8] T. Jormalainen, S. Louhenkilpi: A model for predicting the melt temperature in the ladle and in the tundish as a function of operating parameters during continuous casting, *Steel Research International*, 77 (2006), 472-484.

[9] S. Sonoda, N. Murata, H. Hino, H. Kitada, M. Kano: A statistical model for predicting the liquid steel temperature in ladle and tundish by bootstrap filter, *ISIJ International*, 52 (2012), 1086-1091.

[10] N. Gupta, S. Chandra: Temperature prediction model for controlling casting superheat temperature, *ISIJ International*, 44 (2004), 1517-1526.

[11] T. A. Johansen, B. A. Foss: Representing and learning unmodelled dynamics with neural network memories, *American Control Conference*, (1992), 3037-3043.

[12] D. C. Psichogios, L. H. Ungar: A hybrid neural network first principles approach to process modeling, *AIChE Journal*, 38 (10) (1992), 1499-1511.

[13] B. Li, J. Morris, E. B. Martin: Model selection for partial least squares regression, *Chemometrics and Intelligent Laboratory Systems*, 64 (2002), 79-89

[14] M. Stone: Cross-validatory choice and assessment of statistical predictions, *Journal of the Royal Statistical Society. Series B (Methodological)*, 36 (2) (1974), 111-147

[15] L. Breiman: Random forests, *Machine Learning* 45 (2001), 5-32.

[16] T. K. Ho: Random decision forest, *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Canada, 1 (1995), 278-282.

[17] T. K. Ho: The random subspace method for constructing decision forests, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 (8) (1998), 832-844.

[18] J. L. Bentley: Multidimensional binary search trees used for associative searching, *Communications of the ACM*, 18 (1975), 509-517.

[19] J. H. Friedman, J. Bentely, R. A. Finkel: An algorithm for finding best matches in logarithmic expected time, *ACM Transactions on Mathematical Software*, 3 (3)(1977), 209-226.

[20] N. J. Gordon, D. J. Salmond, A. F. M. Smith: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings*, 140 (1993), 107-113.

Chapter 2

Prediction of Molten Steel Temperature in Tundish

2.1 Introduction

The steel industry faces stiff competition in the global market, and each steel company has to realize stable and efficient operation and produce high quality products satisfying various customer demand. In steel making, tundish is a vessel used for delivering molten steel from a ladle to a mold in the continuous casting process and control of the molten steel temperature in the tundish (TD temp) is one of the key factors to realizing stable operation. If TD temp is too high, breakouts may occur and cause tremendous increase in maintenance cost and productivity loss. When the temperature is too low, clogging in the tundish nozzle occurs, which causes disruptions in the casting process. However, no effective manipulated variable is available after the secondary refining process to control TD temp. To realize the target TD temp, therefore, it is necessary to adjust the molten steel temperature in the secondary refining process (Ruhrstahl-Heraeus degassing process), RH temp. To control TD temp by manipulating RH temp, a model relating TD temp and RH temp needs to be constructed. In the past, vari-

ous models such as first-principle models [1]−[6], statistical models [7], and gray-box models [8] have been proposed to predict molten steel temperature. In this chapter a novel gray-box model is proposed to estimate molten steel temperature in a continuous casting process at a steel making plant by combining a first-principle model and a statistical model. The first-principle model was developed on the basis of computational fluid dynamics (CFD) simulations to simplify the model. Since the derived first-principle model was not able to estimate the molten steel temperature in the tundish with sufficient accuracy, statistical models were developed to estimate the estimation errors of the first-principle model through partial least squares (PLS) and random forests (RF). As a result of comparing the three models, i.e., the first-principle model, the PLS-based gray-box model, and the RF-based gray-box model, the RF-based gray-box model achieved the best estimation performance. The proposed gray-box model was applied to the real process data and the results demonstrated its advantage over other models.

2.2 First-Principle Model

In this section, the proposed first-principle model to estimate TD temp is explained. This first-principle model consists of two parts; the first one models phenomena during the transportation period from the secondary refining to the continuous casting, and the second one models phenomena during the casting period, see **Figure 2.1**. In this model, the degradation of ladles are also taken into account.

2.2.1 First-Principle Model for Transportation Period

Molten Steel in Ladle

It is assumed that the ladle is a cylinder of radius R_{l}. On the basis of the CFD simulation results, indicating that thermal stratification is formed

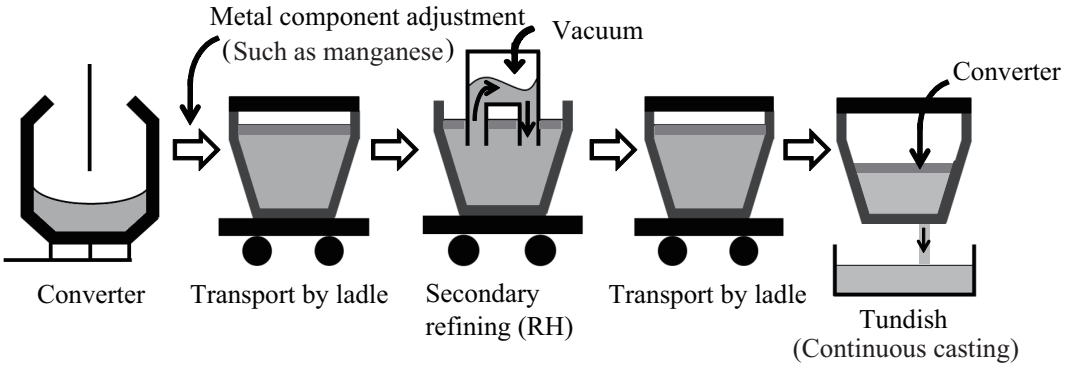


Figure 2.1: Process flow diagram of the steel making process

vertically in the standing ladle due to natural convection [5], the molten steel temperature is modeled as a function of time t and position z from the bottom of the ladle.

$$T_{\mathrm{m}}(z,t) = \overline{T}_{\mathrm{m}}(t) + k(t)\left(\sqrt{\frac{z}{H_{\mathrm{m}}}} - \frac{2}{3}\right) \tag{2.1}$$

where T_{m} is the molten steel temperature, $\overline{T}_{\mathrm{m}}$ is its average, k denotes the difference between the molten steel temperature at the top and the bottom of the ladle, and H_{m} is the depth of the molten steel in the ladle.

The results of CFD simulations have shown that the temperature difference is a function of time [5]; thus it is modeled with parameter α.

$$k(t) = \alpha t \tag{2.2}$$

The initial molten steel temperature is assumed to be homogeneous and the same as RH temp because the molten steel in the ladle is properly stirred. Thus, $k(t) = 0$ at $t = 0$. The transition of average molten steel temperature

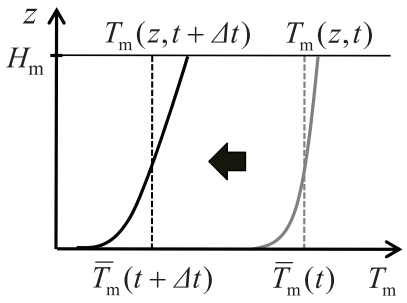


Figure 2.2: Update of molten steel temperature in ladle T_m

$\overline{T}_{\mathrm{m}}(t)$ is calculated through the heat balance equation.

$$\rho_{\mathrm{m}} c_{\mathrm{m}} \pi R_{\mathrm{l}}^2 H_{\mathrm{m}} \frac{d\overline{T}_{\mathrm{m}}(t)}{dt} = -2\pi R_{\mathrm{l}} \int_0^{H_{\mathrm{m}}} U_{\mathrm{w}}(T_{\mathrm{m}}(z,t) - T_{\mathrm{am}}) dz$$
$$-\pi R_{\mathrm{l}}^2 U_{\mathrm{b}}(T_{\mathrm{m}}(0,t) - T_{\mathrm{am}}) - \pi R_{\mathrm{l}}^2 h_1(T_{\mathrm{m}}(H_{\mathrm{m}},t) - T_{\mathrm{sl}}(t)) \qquad (2.3)$$

where ρ_{m} and c_{m} are the density and the heat capacity of the molten steel, respectively. U_{b} and U_{w} are the overall heat transfer coefficients of the ladle bottom and the ladle wall, respectively. T_{am} and T_{sl} are the ambient temperature and the slag temperature, respectively. In addition, h_1 denotes the heat transfer coefficient between the molten steel and the slag. The left side of Eq. (2.3) represents the time change of the molten steel enthalpy. The first, second, and third terms of the right side represent the heat conduction from the molten steel to the ladle wall, to the ladle bottom, and to the slag, respectively.

The average molten steel temperature is gradually decreased with time by Eq. (2.3). The temperature profile in the ladle at two different moments is shown in **Figure 2.2**.

Slag in Ladle

Slag in the standing ladle, which is generated in the convertor, keeps the molten steel at high temperature. The heat balance of the slag is modeled as a lumped parameter system.

$$\rho_{\mathrm{sl}} c_{\mathrm{sl}} \pi R_{\mathrm{l}}^2 H_{\mathrm{sl}} \frac{dT_{\mathrm{sl}}(t)}{dt} = \pi R_{\mathrm{l}}^2 h_1 (T_{\mathrm{m}}(H_{\mathrm{m}}, t) - T_{\mathrm{sl}}(t)) - \pi R_{\mathrm{l}}^2 \varepsilon_{\mathrm{sl}} \sigma (T_{\mathrm{sl}}(t)^4 - T_{\mathrm{am}}^4) - \pi R_{\mathrm{l}}^2 h_2 (T_{\mathrm{sl}}(t) - T_{\mathrm{a1}}) - 2\pi R_{\mathrm{l}} H_{\mathrm{sl}} U_{\mathrm{w}} (T_{\mathrm{sl}}(t) - T_{\mathrm{am}}) \tag{2.4}$$

where ρ_{sl}, c_{sl}, and $\varepsilon_{\mathrm{sl}}$ are the density, the heat capacity, and the emissivity of the slag, respectively. H_{sl} denotes the slag layer thickness, h_2 the heat transfer coefficient between the slag and the air in the ladle, T_{a1} the air temperature in the ladle, and σ the Stefan-Boltzmann coefficient. The left side of Eq. (2.4) represents the time change of the slag enthalpy. The first, second, third and fourth terms of the right side represent the heat conduction from the molten steel to the slag, the radiation from the slag to the wall of the ladle, the heat conduction from the slag to the air in the ladle, and that from the slag to the ladle wall, respectively. The ladle wall temperature, which affects the radiation, is assumed to be equal to the air temperature.

Ladle Degradation

Due to the repeated use of the ladle, the walls of the ladle gradually degrade. The effect of ladle degradation on the heat conduction flux from the molten steel to the external environment has been discussed in the literature. One study [9] describes the factors which cause degradation of ladle while another study [10] develops a CFD model to relate the heat losses from ladle with the reduction in ladle walls and bottom thickness. To avoid computational complexity and build a simple model, it is assumed that the overall heat

transfer coefficients gradually increase with the number of repeated usage, u. Furthermore, the ratio of increase of the overall heat transfer coefficient of the ladle wall is the same as that of the ladle bottom. In addition, it is assumed that the temperature difference between the top and the bottom of ladle increases with the increase of u. The relations are expressed by

$$U_{\mathrm{w}}(u) = \beta U_{\mathrm{b}}(u) \tag{2.5}$$

$$U_{\mathrm{b}}(u) = U_{\mathrm{b0}} + \gamma\sqrt{u} \tag{2.6}$$

$$\alpha(u) = \alpha_0 + \eta\sqrt{u} \tag{2.7}$$

where β, U_{b0}, γ, α_0 and η are constants.

2.2.2 First-Principle Model for Casting Period

Molten Steel in Ladle

It is assumed that the volumetric flow Q from the ladle to the tundish is constant. Let $T_{\mathrm{md}}(z,t)$ be the temperature profile in the ladle during the discharging period, and the starting moment of discharging regarded as $t = 0$. Then, the enthalpy balance of the molten steel in the ladle during the withdrawal period is given by the following equations.

$$\rho_{\mathrm{m}} c_{\mathrm{m}} \pi R_i^2 H_{\mathrm{m}} \frac{d\overline{T}_{\mathrm{md}}(t)}{dt} = -2\pi R_i \int_0^{H_{\mathrm{m}}} U_{\mathrm{w}}(T_{\mathrm{md}}(z,t) - T_{\mathrm{am}})dz$$
$$-\pi R_i^2 U_{\mathrm{b}}(T_{\mathrm{md}}(0,t) - T_{\mathrm{am}}) - \pi R_i^2 h_1(T_{\mathrm{md}}(H_{\mathrm{m}}t) - T_{\mathrm{sl}}(t))$$
$$-\rho_{\mathrm{m}} c_{\mathrm{m}} Q(T_{\mathrm{md}}(0,t) - \overline{T}_{\mathrm{m}}(t)) \tag{2.8}$$

$$H_{\mathrm{m}}(t) = H_{\mathrm{m}}(0) - \frac{Q}{\pi R^2} t \tag{2.9}$$

$$T_{\mathrm{md}}(z,0) = T_{\mathrm{m}}(z,t_f) \tag{2.10}$$

where t_{f} is the ending moment of the transportation period. The last term of the right hand side of Eq. (2.8) shows the effect of withdrawal. H_{m} is gradually decreased by the withdrawal of the molten steel. It is assumed that the shape of the temperature profile in the ladle is not changed during the discharging period, i.e., the temperature profile during the discharging period is assumed to be given by the following equation:

$$T_{\mathrm{md}}(z,t) = \overline{T}_{\mathrm{md}}(t) + T_{\mathrm{md}}(z + \frac{Q}{\pi R^2}t, 0) - \frac{1}{H_{\mathrm{m}}(t)} \int_0^{H_{\mathrm{m}}(t)} T_{\mathrm{md}}(z + \frac{R}{\pi R^2}t, 0)dz \tag{2.11}$$

The last term of Eq. (2.11) is added so that the average temperature of T_{m} becomes $\overline{T}_{\mathrm{m}}$. An example of temperature profiles at the starting moment and the middle of the discharging period is shown in **Figure 2.3**. For the enthalpy balance of the slag layer, Eq. (2.4) is used.

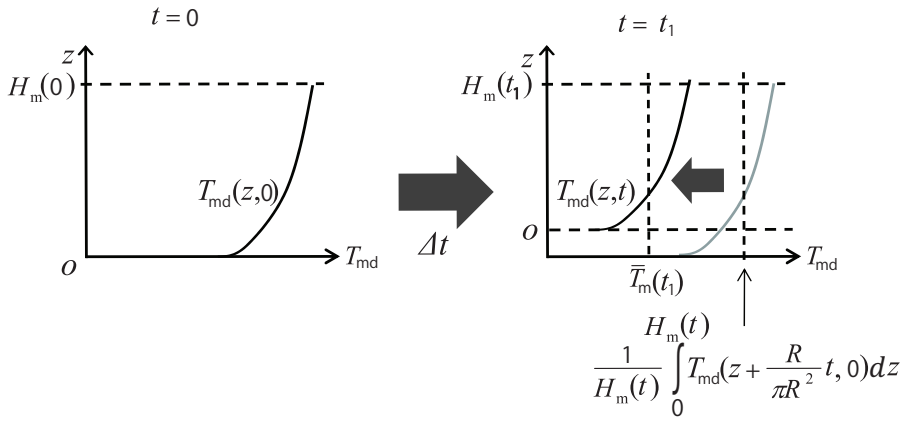


Figure 2.3: Model of molten steel in ladle during the casting period

Molten Steel in the Tundish

It is assumed that the inflow to the tundish is equal to the outflow from the tundish and also the depth of the molten steel in the tundish is constant. The CFD simulations have indicated that temperature in the tundish is distributed in the flow direction [11]. Thus, the tundish is modeled as a compartment model consisting of N_{t} isothermal baths connected in series as shown in **Figure 2.4**. The heat balance of the kth bath is expressed by Eqs. (2.12) and (2.13)

$$\rho_{\mathrm{m}} c_{\mathrm{m}} W H \frac{L}{N_{\mathrm{t}}} \frac{dT_{\mathrm{t}}^{(k)}(t)}{dt} = \rho_{\mathrm{m}} c_{\mathrm{m}} Q T_{\mathrm{t}}^{(k-1)}(t) - \rho_{\mathrm{m}} c_{\mathrm{m}} Q T_{\mathrm{t}}^{(k)}(t) - S_{\mathrm{t}} U_{\mathrm{t}} (T_{\mathrm{t}}^{(k)}(t) - T_{\mathrm{am}}) - W \frac{L}{N_{\mathrm{t}}} \varepsilon_{\mathrm{t}} \sigma (T_{\mathrm{t}}^{(k)}(t)^4 - T_{\mathrm{am}}^4) - W \frac{L}{N_{\mathrm{t}}} h_3 (T_{\mathrm{t}}^{(k)}(t) - T_{\mathrm{a2}}) \tag{2.12}$$

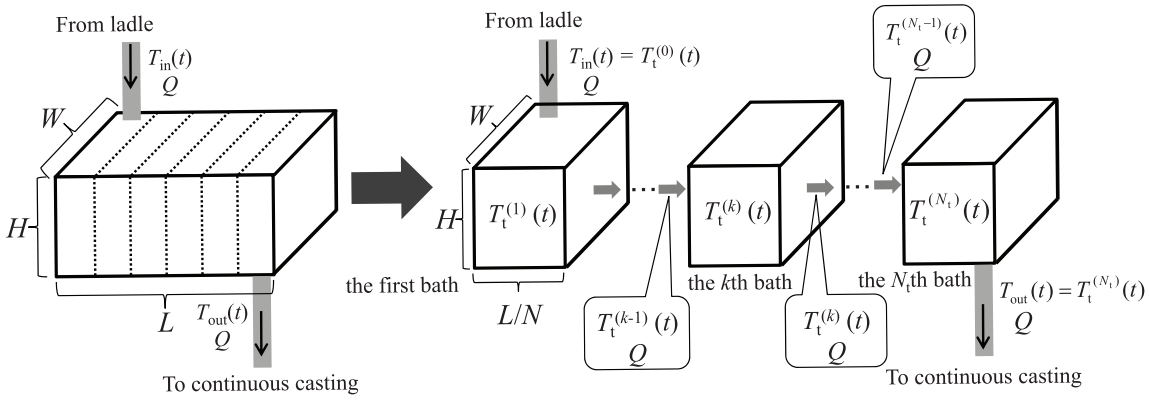


Figure 2.4: Compartment model of molten steel in tundish

$$S_t = \begin{cases} W\dfrac{L}{N_\mathrm{t}} + 2H\dfrac{L}{N_\mathrm{t}} + WH & (k = 1, N_\mathrm{t}) \\ W\dfrac{L}{N_\mathrm{t}} + 2H\dfrac{L}{N_\mathrm{t}} & (k = 2, 3, \cdots, N_\mathrm{t} - 1) \end{cases} \tag{2.13}$$

where W, H, and L denote the width, the height, and the length of molten steel in the tundish, respectively. $T_\mathrm{t}^{(k)}$ is the tundish temperature in the kth bath. S_t denotes the contact area between the molten steel and the tundish, U_t the overall heat transfer coefficient of the tundish, ε_t the emissivity of the molten steel, T_a2 the air temperature in the tundish, and h_3 the heat transfer coefficient between the molten steel and the air. $T_\mathrm{t}^{(0)}(t)$ is equal to $T_\mathrm{in}(t)$ because the molten steel poured from the ladle flows into the first bath. The left side of Eq. (2.12) represents the time change of the molten steel enthalpy. The first to fifth terms of the right side represent the inflow enthalpy, the outflow enthalpy, the heat conduction from the molten steel to the tundish wall, the radiation from the molten steel to the tundish wall and the heat conduction from the molten steel to the air in the tundish, respectively. The tundish wall temperature is assumed to be equal to the air temperature T_am, which is assumed to be constant.

2.2.3 Parameter Fitting

The physical model contains 13 parameters to be identified: α_0, γ, η, h_1, h_2, h_3, $U_{\mathrm{b}0}$, U_{t}, $T_{\mathrm{a}1}$, $T_{\mathrm{a}2}$, $\varepsilon_{\mathrm{sl}}$, ε_{t} and β. The first ten parameters were identified through the least squares algorithm using real process data and $\varepsilon_{\mathrm{sl}}$, ε_{t} and β are given in advance from the experience of engineers. The dimensions of the ladle and the tundish, and the physical properties of steel are also given in advance. N_{t} is set equal to three based on the CFD study by Odenthal et al. [11]. The input variables of the physical model are the number of ladle usage, the weight of the molten steel in the ladle, the initial temperature of steel in the ladle, the transportation time, the casting flow rate, and the initial temperature of each bath composing the tundish. The output variable is the temperature of the last bath, TD temp at a certain point of time. A total of 1155 samples were used for parameter estimation.

The developed model was validated by using 289 samples, which were used only for validation. The prediction result through the first-principle model is shown in **Figure 2.5**. The prediction performance was evaluated on the basis of the root-mean-square error (RMSE) and the correlation coefficient between scaled reference (measured values) and predicted values (r). RMSE of scaled TD temp was 2.72, which did not meet our specification. Hence, the prediction performance of the developed first-principle model was not sufficient for its industrial application.

2.3 Gray-box Model

To improve the prediction performance, a gray-box model is constructed by combining the first-principle model and a statistical model sequentially; the statistical model is built so as to compensate the error of the first-principle model. This type of gray-box model, called the parallel gray-box model, is developed through the following steps.

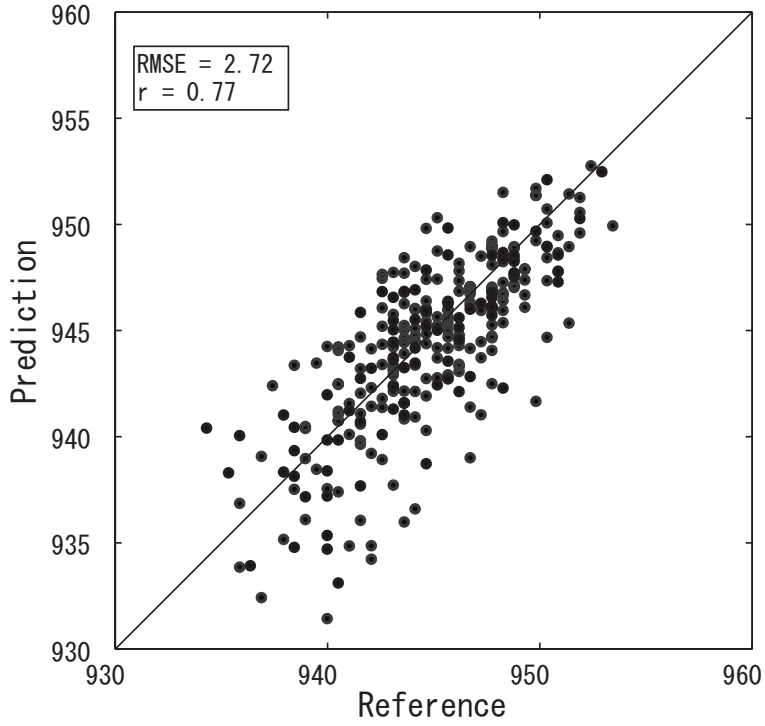


Figure 2.5: Prediction of TD temp (scaled) through the first-principle model

i. Build a first-principle model f_{fp} to predict an output variable y from input variables X_{fp}.

$$\hat{y}_{\mathrm{fp}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \boldsymbol{\theta}) \tag{2.14}$$

where $\hat{y}_{\mathrm{fp}}$ is the prediction of y and $\boldsymbol{\theta}$ is a parameter vector.

ii. Estimate the parameters $\boldsymbol{\theta}$ by solving the optimization problem of minimizing the sum of squared errors.

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{n=1}^{N_\mathrm{d}} e_{\mathrm{fp},n}^2 \tag{2.15}$$

$$e_{\mathrm{fp},n} = y_n - f_\mathrm{fp}(X_{\mathrm{fp},n}, \boldsymbol{\theta}) \tag{2.16}$$

$$\boldsymbol{\theta}_\mathrm{L} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_\mathrm{U} \tag{2.17}$$

where $X_{\mathrm{fp},n}$ and y_n are the nth sample of input and output variables, respectively. N_d is the number of samples used for developing the model. $\boldsymbol{\theta}_\mathrm{L}$ and $\boldsymbol{\theta}_\mathrm{U}$ are lower and upper bound vectors of parameters which are determined in advance.

iii. Build a statistical model f_pa to predict the output error e_fp from input variables X by solving the following optimization problem.

$$\tilde{\boldsymbol{\varphi}}_\mathrm{pa} = \underset{\boldsymbol{\varphi}_\mathrm{pa}}{\operatorname{argmin}} \sum_{n=1}^{N_\mathrm{d}} (e_{\mathrm{fp},n} - f_\mathrm{out}(X_n, \boldsymbol{\varphi}_\mathrm{pa}))^2 \tag{2.18}$$

$$\hat{e}_{\mathrm{fp},n} = f_\mathrm{pa}(X_n, \boldsymbol{\varphi}_\mathrm{pa}) \tag{2.19}$$

where $\boldsymbol{\varphi}_\mathrm{pa}$ is parameters in the statistical model. In general, X_fp is a subset of X.

iv. Build a gray-box model by combining the first-principle model and the statistical model.

Table 2.1: Prediction results of five models

Modeling method	RMSE	r
First-principle model	2.72	0.77
Statistical model (PLS)	2.20	0.82
Statistical model (RF)	2.09	0.85
Gray-box model (PLS)	2.16	0.83
Gray-box model (RF)	1.73	0.89

$$\hat{y}_{\mathrm{pa}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \tilde{\boldsymbol{\theta}}) + f_{\mathrm{pa}}(X, \tilde{\boldsymbol{\varphi}}_{\mathrm{pa}}) \tag{2.20}$$

where $\hat{y}_{\mathrm{pa}}$ is the prediction of y by using the parallel gray-box model.

2.3.1 Model Validation

Finally, statistical models and gray-box models were constructed and their prediction performance was compared by using real process data provided by Sumitomo Metal Industries, Ltd.

The number of samples was 1444; 1155 samples (80 %) were used for modeling and the other 289 samples (20 %) were used for validation. The prediction results of five models are summarized in **Table 2.1**.

In addition, **Figure 2.6** shows the prediction results through the statistical models, i.e., PLS and RF, and the gray-box models that integrate the first-principle model with PLS and RF.

The input variables of the statistical models were a total of 51 measured variables in the processes from the converter to the tundish.

The validation results show that the proposed gray-box model combining the first-principle model and the RF model achieved the highest prediction accuracy. Its RMSE was 17 % and 36 % smaller than that of RF and the

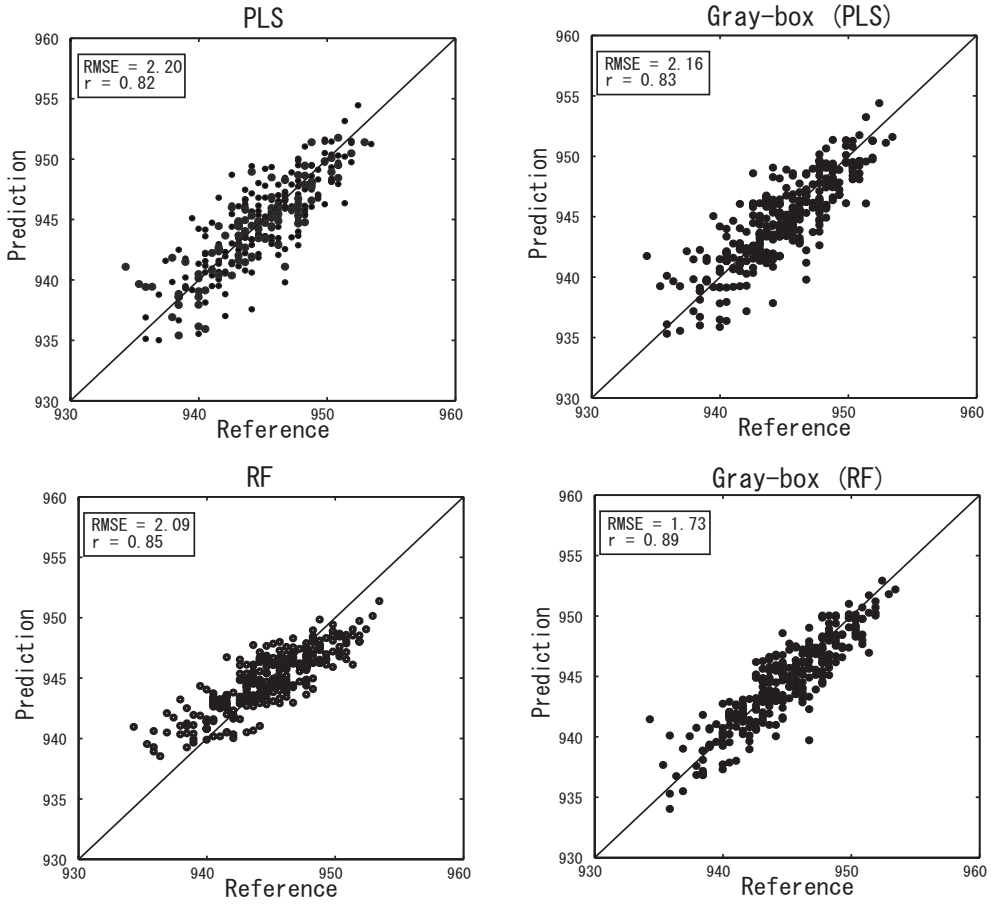


Figure 2.6: Prediction of TD temp through PLS model, RF model, PLS-based gray-box model, and RF-based gray-box model

first-principle model, respectively. The results have clarified the advantage of the proposed gray-box model over other models. It is important that the prediction performance of the RF model alone was better than the gray-box model, in which the PLS model was used to estimate the prediction errors of the first-principle model. This shows that PLS lacks the capability to extract enough information from the data which was needed to overcome the errors of the firs-principle model. Although RF can build a nonlinear process model, its direct application is not always the best approach.

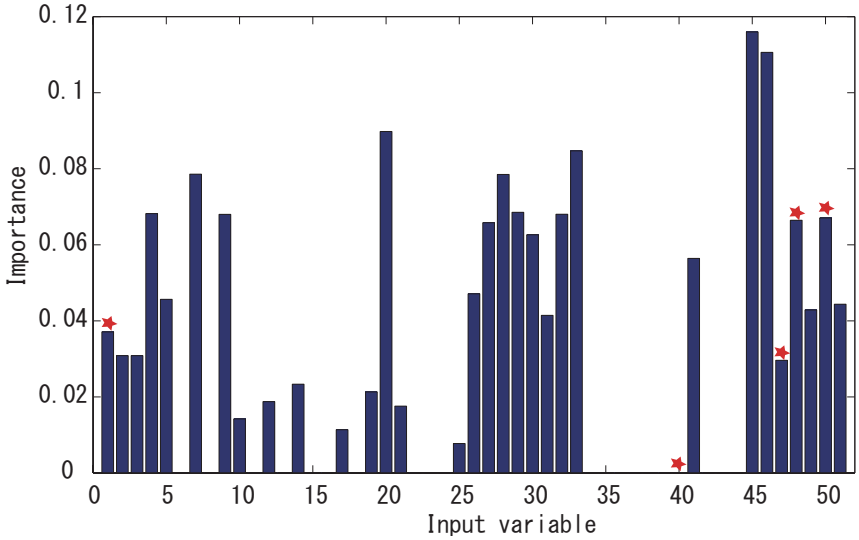


Figure 2.7: Importance of input variables for RF model in gray-box model

RF has an ability to calculate the "importance" of each input variable to the output variables. This "importance" of input variables is determined on the basis of increase in prediction error of out-of-bag data (OOB) by the corresponding trees in the RF, in the absence of that particular input variable. By utilizing this feature of RF, important input variables of the RF model integrated with the first-principle model were identified as shown in **Figure 2.7**.

The important variables, which contributed toward estimating the prediction error of the first-principle models, were the 20th, 33rd, 45th and 46th. The 20th variable is the amount of alloy added to the converter, 33rd is the inlet temperature of the RH, 45th is the residence time of RH and 46th is the time duration from start of RH operation to the RH temp measurement. In other words, these four variables are crucial in describing the phenomena that cannot be modeled by the first-principle model. The variables used in the first-principle model are marked by red star ($\star$) in the Figure 2.7. These variables except the RH temp, i.e., the 40th, still have some valuable

information which are utilized by the RF.

2.4 Conclusions

In this study, a new gray-box model to predict the molten steel temperature in tundish (TD temp) in a steel making plant was proposed and was applied to the real process data. Partial least squares (PLS) and random forests (RF) were used to build statistical models. The results have clearly shown the advantage of the developed RF based gray-box model over the first-principle model and the statistical models.

Bibliography

[1] A. Zoryk, P. M. Reid: On-line liquid steel temperature control, *Iron and Steelmaker*, 20 (1993), 21-27.

[2] A. G. Belkovskii, Y. L. Kats: Mathematical model of the cooling of steel in a small ladle, *Metallurgist*, 53 (2009), 261-273.

[3] P. R. Austin, J. M. Camplin: J. Herbertson, I. J. Taggart, Mathematical modeling of thermal stratification and drainage of steel ladles, *ISIJ International*, 32 (2) (1992), 196-202.

[4] J. L. Xia, T. Ahokainen: Transient flow and heat transfer in a steelmaking ladle during the holding period, *Metallurgical and Materials Transactions B*, 32 (2001), 733-741.

[5] J. R. S. Zabadal, M. T. M. B. Vilhena, S. Q. B. Leit: Heat transfer process simulation by finite differences for online control of ladle furnaces, *Ironmaking and Steelmaking*, 31 (2004), 227-232.

[6] T. Jormalainen, S. Louhenkilpi: A model for predicting the melt temperature in the ladle and in the tundish as a function of operating parameters during continuous casting, *Steel Research International*, 77 (2006), 472-484.

[7] S. Sonoda, N. Murata, H. Hino, H. Kitada, M. Kano: A statistical model for predicting the liquid steel temperature in ladle and tundish by bootstrap filter, *ISIJ International*, 52 (2012), 1086-1091.

[8] N. Gupta, S. Chandra: Temperature prediction model for controlling casting superheat temperature, *ISIJ International*, 44 (2004), 1517-1526.

[9] T. P. Fredman: Heat transfer in steelmaking ladle refractories and steel temperature, *Scandinavian Journal of Metallurgy*, 29 (2000), 232-258.

[10] A. Tripathi, J. K. Saha, J. B. Singh, S. K. Ajmani: Numerical simulation of heat transfer phenomenon in steel making ladle, *ISIJ International*, 52 (9) (2012), 1591-1600.

[11] H. J. Odenthal, M. Javurek, M. Kirschen, N. Vogl: CFD benchmark for a single strand tundish, *Steel Research International*, 81 (7) (2010), 529-541.

Chapter 3

Control of Molten Steel Temperature in Tundish

3.1 Introduction

In this chapter another model is proposed to improve the prediction accuracy of the parallel gray-box model developed in chapter 2, and make it able to control molten steel temperature in tundish. The parallel gray-box model developed in chapter 2 achieved high prediction performance by combining a first-principle model and a statistical model, however the accuracy of the first-principle model is still important. The first-principle model have various parameters which were determined by using data and kept constant because it was difficult to find a relationship between the parameters and the operating conditions. The parallel gray-box model can extract information from data that is not used in the first-principle model and also it can overcome the limitations imposed by the structure of the first-principle model. However, it does not investigate the outdated parameters of the first-principle model and simply compensate the errors, by the statistical models. The outdated parameters should be searched and updated according to the changing operation condition. This chapter presents a new gray-box model

that can overcome such deficiency and can predict molten steel temperature with high accuracy. To achieve this goal, a parameter in the first-principle model is estimated from process variables with a nonlinear statistical model. In addition, another statistical model is used to compensate the prediction errors. Random forests (RF) is adopted in this work to build statistical models. To use the proposed model for control of the TD temp, it is used in inverse format with the help of bisection method. The inverse approach is adopted because the continuous casting process has no manipulated variable to directly control TD temp, and therefore the proposed gray-box model is used to determine the molten steel temperature in the Ruhrstahl-Heraeus degassing process (RH degasser) at the end of its operation from the target TD temp. The proposed modeling and control strategy is validated through its application to real operation data at a steel work.

3.2 Gray-box Models

In this section, gray-box models are categorized into three types, parallel gray-box models [1], serial gray-box models [2], and combined gray-box models. A generalized framework of these gray-box models is explained blow:

3.2.1 Parallel Gray-box Model

For comparison, the parallel gray-box model explained at section 2.3 is shown again in this subsection. It is developed through the following steps.

i. Build a first-principle model f_{fp} to predict an output variable y from input variables X_{fp}.

$$\hat{y}_{\text{fp}} = f_{\text{fp}}(X_{\text{fp}}, \boldsymbol{\theta}) \tag{3.1}$$

where $\hat{y}_{\mathrm{fp}}$ is the prediction of y and $\boldsymbol{\theta}$ is a parameter vector.

ii. Estimate the parameters $\boldsymbol{\theta}$ by solving the optimization problem of minimizing the sum of squared errors.

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{n=1}^{N_{\mathrm{d}}} e_{\mathrm{fp},n}^2 \tag{3.2}$$

$$e_{\mathrm{fp},n} = y_n - f_{\mathrm{fp}}(X_{\mathrm{fp},n}, \boldsymbol{\theta}) \tag{3.3}$$

$$\boldsymbol{\theta}_{\mathrm{L}} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\mathrm{U}} \tag{3.4}$$

where $X_{\mathrm{fp},n}$ and y_n are the nth sample of input and output variables, respectively. N_{d} is the number of samples used for developing the model. $\boldsymbol{\theta}_{\mathrm{L}}$ and $\boldsymbol{\theta}_{\mathrm{U}}$ are lower and upper bound vectors of parameters which are determined in advance.

iii. Build a statistical model f_{pa} to predict the output error e_{fp} from input variables X by solving the following optimization problem.

$$\tilde{\boldsymbol{\varphi}}_{\mathrm{pa}} = \underset{\boldsymbol{\varphi}_{\mathrm{pa}}}{\operatorname{argmin}} \sum_{n=1}^{N_{\mathrm{d}}} (e_{\mathrm{fp},n} - f_{\mathrm{pa}}(X_n, \boldsymbol{\varphi}_{\mathrm{pa}}))^2 \tag{3.5}$$

$$\hat{e}_{\mathrm{fp},n} = f_{\mathrm{pa}}(X_n, \boldsymbol{\varphi}_{\mathrm{pa}}) \tag{3.6}$$

where $\boldsymbol{\varphi}_{\mathrm{pa}}$ is parameters in the statistical model. In general, X_{fp} is a subset of X.

iv. Build a gray-box model by combining the first-principle model and the statistical model.

$$\hat{y}_{\mathrm{pa}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \tilde{\boldsymbol{\theta}}) + f_{\mathrm{pa}}(X, \tilde{\boldsymbol{\varphi}}_{\mathrm{pa}}) \tag{3.7}$$

where $\hat{y}_{\mathrm{pa}}$ is the prediction of y by using the parallel gray-box model.

3.2.2 Serial Gray-box Model

In the parallel gray-box model, the optimal parameters $\tilde{\boldsymbol{\theta}}$ are constant, although some parameters depend on the operating conditions. This simplification may deteriorate the prediction performance of the first-principle model. Thus, the serial gray-box model is used to estimate the parameters as functions of input variables. The general procedure of developing the general structure of the serial gray-box model is explained in chapter 1. However, in that general procedure it is assumed that the parameter to be modeled by a statistical function is given in advance. Here we propose, a new serial gray-box with a systematic search for an optimum parameter which needs to be updated. The new serial gray-box is constructed by the following steps.

i. A first-principle model $\hat{y}_{\mathrm{fp}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \boldsymbol{\theta})$ is developed.

ii. The parameters $\boldsymbol{\theta}$ in the first-principle model are optimized by Eq. (3.2).

iii. A parameter θ_i is selected, and is optimized for each modeling sample.

$$\tilde{\theta}_{i,n} = \underset{\theta_{i,n}}{\operatorname{argmin}}(y_n - f_{\mathrm{fp}}(X_{\mathrm{fp},n}, \tilde{\boldsymbol{\theta}}_i^c, \theta_{i,n}))^2 \ (n = 1, 2, \cdots, N_{\mathrm{d}}), \quad (i = 1, 2, \cdots, N_{\mathrm{p}}) \tag{3.8}$$

$$\theta_{\mathrm{L},i} \leq \theta_{i,n} \leq \theta_{\mathrm{U},i} \tag{3.9}$$

where $\tilde{\theta}_{i,n}$ is the optimal value of the parameter θ_i for sample n and $\tilde{\boldsymbol{\theta}}_i^c$ is the constant vector consisting of the parameters except θ_i. N_{p} is the number of parameters in the first-principle model. $\theta_{\mathrm{L},i}$ and $\theta_{\mathrm{U},i}$ are lower and upper bounds of the ith parameter which are determined in advance.

iv. The step iii is repeated for all the parameters one by one, and the parameter which achieves the smallest values of $(y_n - f_{\mathrm{fp}}(X_{\mathrm{fp},n}, \tilde{\boldsymbol{\theta}}_i^c, \theta_{i,n}))^2$ is selected.

v. A statistical model f_{se} is constructed to estimate the selected parameter θ_i from the input variables X by solving the following optimization problem.

$$\tilde{\boldsymbol{\varphi}}_{\mathrm{se},i} = \underset{\boldsymbol{\varphi}_{\mathrm{se},i}}{\operatorname{argmin}} \sum_{n=1}^{N_{\mathrm{d}}} (\tilde{\theta}_{i,n} - f_{\mathrm{se}}(X_n, \boldsymbol{\varphi}_{\mathrm{se},i}))^2 \tag{3.10}$$

where $\tilde{\boldsymbol{\varphi}}_{\mathrm{se},i}$ is the optimal values of parameters.

vi. By combining the first-principle model and the statistical model, a gray-box model is derived:

$$\hat{y}_{\text{se}} = f_{\text{fp}}(X_{\text{fp}}, \tilde{\boldsymbol{\theta}}_i^c, \hat{\tilde{\theta}}_i) \tag{3.11}$$

$$\hat{\tilde{\theta}}_i = f_{\text{se}}(X, \tilde{\boldsymbol{\varphi}}_{\text{se},i}) \tag{3.12}$$

The serial gray-box modeling leads to better understanding of the process by recursively searching the most effective parameters to minimize the effect of changing operating conditions on the prediction performance of the first-principle model. On contrary, the parallel gray-box modeling does not investigate the outdated parameters of the first-principle model and simply compensate the errors, by the statistical models. Although, the serial gray-box modeling approach update the selected parameters and minimize the prediction errors better than the parallel gray-box modeling, there is still a significant amount of errors left. To overcome this remaining errors, another approach called the combined gray-box modeling is adopted.

3.2.3 Combined Gray-box Model

The combined gray-box model is generated by embedding a parallel gray-box model with a serial gray-box model. In the combined gray box model, the serial gray-box model is used as the first-principle model of the parallel gray-box model. The combined gray-box model is constructed by the following steps.

i. Build a serial gray-box model to predict an output variable y.

$$\hat{y}_{\text{se}} = f_{\text{fp}}(X_{\text{fp}}, \tilde{\boldsymbol{\theta}}^c, \hat{\tilde{\theta}}) \tag{3.13}$$

$$e_{\mathrm{se}} = y - \hat{y}_{\mathrm{se}} \tag{3.14}$$

where $\hat{y}_{\mathrm{se}}$ is the prediction of y and e_{se} is the output errors of the serial gray-box model.

ii. Build a statistical model f_{com} to predict the output errors e_{se} from input variables X by solving the following optimization problem.

$$\tilde{\boldsymbol{\varphi}}_{\mathrm{com}} = \underset{\boldsymbol{\varphi}_{\mathrm{com}}}{\operatorname{argmin}} \sum_{n=1}^{N_{\mathrm{d}}} (e_{\mathrm{se},n} - f_{\mathrm{com}}(X_n, \boldsymbol{\varphi}_{\mathrm{com}}))^2 \tag{3.15}$$

where $\tilde{\boldsymbol{\varphi}}_{\mathrm{com}}$ is the optimized parameters in the statistical model.

iii. Build a combined gray-box model by combining the serial gray-box model and the statistical model.

$$\hat{y}_{\mathrm{com}} = f_{\mathrm{fp}}(X_{\mathrm{fp}}, \tilde{\boldsymbol{\theta}}^c, \hat{\tilde{\theta}}) + f_{\mathrm{com}}(X, \tilde{\boldsymbol{\varphi}}_{\mathrm{com}}) \tag{3.16}$$

where $\hat{y}_{\mathrm{com}}$ is the prediction of y by using the combined gray-box model.

3.3 First-Principle Model

First-principle model in this chapter is the same as that in chapter 2 [3]. In order to make the location of the parameters clearer, dominant equations are explained again.

3.3.1 First-Principle Model for Transportation Period

The molten steel temperature in a ladle is modeled as a function of time t and position z from the bottom of the ladle.

$$T_{\mathrm{m}}(z,t) = \overline{T}_{\mathrm{m}}(t) + k(t)\left(\sqrt{\frac{z}{H_{\mathrm{m}}}} - \frac{2}{3}\right) \tag{3.17}$$

$$k(t) = \alpha t \tag{3.18}$$

where T_{m} is the molten steel temperature, $\overline{T}_{\mathrm{m}}$ is its average, k denotes the difference between the molten steel temperature at the top and the bottom of the ladle, and H_{m} is the depth of the molten steel in the ladle.

The initial molten steel temperature is assumed to be homogeneous and the same as RH temp because the molten steel in the ladle is properly stirred. Thus, $k(t) = 0$ at $t = 0$. The transition of average molten steel temperature $\overline{T}_{\mathrm{m}}(t)$ is calculated through the heat balance equation.

$$\begin{aligned}\rho_{\mathrm{m}} c_{\mathrm{m}} \pi R_{\mathrm{l}}^2 H_{\mathrm{m}} \frac{d\overline{T}_{\mathrm{m}}(t)}{dt} = -2\pi R_{\mathrm{l}} \int_0^{H_{\mathrm{m}}} U_{\mathrm{w}}(T_{\mathrm{m}}(z,t) \\ -T_{\mathrm{am}})dz - \pi R_{\mathrm{l}}^2 U_{\mathrm{b}}(T_{\mathrm{m}}(0,t) - T_{\mathrm{am}}) \\ -\pi R_{\mathrm{l}}^2 h_1(T_{\mathrm{m}}(H_{\mathrm{m}},t) - T_{\mathrm{sl}}(t))\end{aligned} \tag{3.19}$$

where ρ_{m} and c_{m} are the density and the heat capacity of the molten steel, respectively. U_{b} and U_{w} are the overall heat transfer coefficients of the ladle bottom and the ladle wall, respectively. T_{am} and T_{sl} are the ambient temperature and the slag temperature, respectively. In addition, h_1 denotes the heat transfer coefficient between the molten steel and the slag.

The heat balance of the slag is modeled as a lumped parameter system.

$$\rho_{\rm sl}c_{\rm sl}\pi R_{\rm l}^2 H_{\rm sl}\frac{dT_{\rm sl}(t)}{dt} = \pi R_{\rm l}^2 h_1(T_{\rm m}(H_{\rm m},t) - T_{\rm sl}(t)) \\ -\pi R_{\rm l}^2\varepsilon_{\rm sl}\sigma(T_{\rm sl}(t)^4 - T_{\rm a1}^4) - \pi R_{\rm l}^2 h_2(T_{\rm sl}(t) - T_{\rm a1}) \\ -2\pi R_{\rm l}H_{\rm sl}U_{\rm w}(T_{\rm sl}(t) - T_{\rm am}) \tag{3.20}$$

where $\rho_{\rm sl}$, $c_{\rm sl}$, and $\varepsilon_{\rm sl}$ are the density, the heat capacity, and the emissivity of the slag, respectively. $H_{\rm sl}$ denotes the slag layer thickness, h_2 the heat transfer coefficient between the slag and the air in the ladle, $T_{\rm a1}$ the air temperature in the ladle, and σ the Stefan-Boltzmann coefficient.

The effect of ladle degradation on the heat conduction flux from the molten steel to the external environment is expressed by the following equations.

$$U_{\rm w}(u) = \beta U_{\rm b}(u) \tag{3.21}$$

$$U_{\rm b}(u) = U_{\rm b0} + \gamma\sqrt{u} \tag{3.22}$$

$$\alpha(u) = \alpha_0 + \eta\sqrt{u} \tag{3.23}$$

where β, $U_{\rm b0}$, γ, α_0 and η are constants.

3.3.2 First-Principle Model for Casting Period

The enthalpy balance of the molten steel in the ladle during the withdrawal period is given by the following equations.

$$\rho_{\rm m}c_{\rm md}\pi R_l^2 H_{\rm m}\frac{d\overline{T}_{\rm m}(t)}{dt} = -2\pi R_l\int_0^{H_{\rm m}} U_{\rm w}(T_{\rm md}(z,t) \\ -T_{\rm am})dz - \pi R_l^2 U_{\rm b}(T_{\rm md}(0,t) - T_{\rm am}) - \pi R_l^2 h_1(T_{\rm md}(H_{\rm m}t) \\ -T_{\rm sl}(t)) - \rho_{\rm m}c_{\rm m}Q(T_{\rm md}(0,t) - \overline{T}_{\rm m}(t)) \tag{3.24}$$

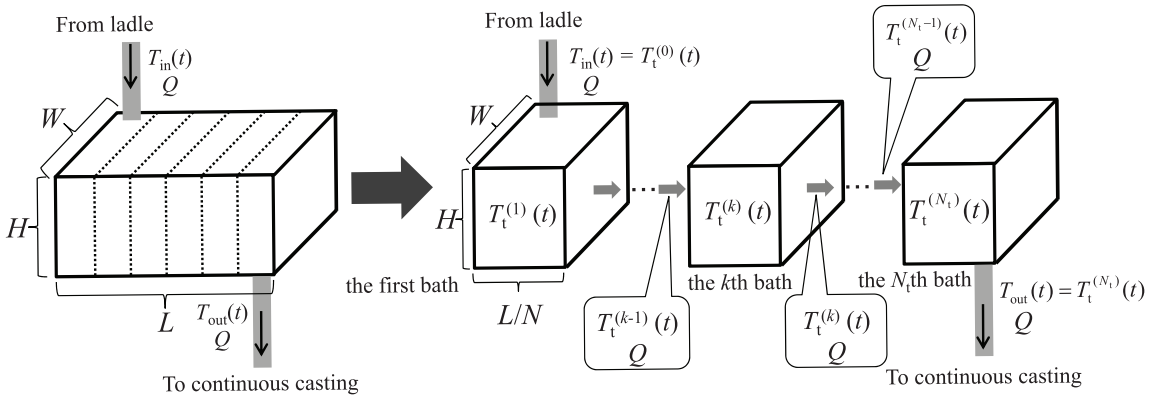


Figure 3.1: Compartment model of molten steel in tundish (Okura et al. [3])

$$H_{\mathrm{m}}(t) = H_{\mathrm{m}}(0) - \frac{Q}{\pi R^2} t \tag{3.25}$$

$$T_{\mathrm{md}}(z, 0) = T_{\mathrm{m}}(z, t_f) \tag{3.26}$$

where ρ_{m} and c_{m} are the density and the heat capacity of the molten steel, respectively. U_{b} and U_{w} are the overall heat transfer coefficients of the ladle bottom and the ladle wall, respectively.

The tundish is modeled as a compartment model consisting of N_{t} isothermal baths connected in series as shown in **Figure 3.1**. The heat balance of the kth bath is expressed by Eqs. (3.27).

$$\begin{aligned} \rho_{\mathrm{m}} c_{\mathrm{m}} W H \frac{L}{N_{\mathrm{t}}} \frac{dT_{\mathrm{t}}^{(k)}(t)}{dt} &= \rho_{\mathrm{m}} c_{\mathrm{m}} Q T_{\mathrm{t}}^{(k-1)}(t) - \rho_{\mathrm{m}} c_{\mathrm{m}} Q T_{\mathrm{t}}^{(k)}(t) \\ &\quad - S_{\mathrm{t}}^{(k)} U_{\mathrm{t}} (T_{\mathrm{t}}^{(k)}(t) - T_{\mathrm{am}}) - W \frac{L}{N_{\mathrm{t}}} \varepsilon_{\mathrm{t}} \sigma (T_{\mathrm{t}}^{(k)}(t)^4 \\ &\quad - T_{\mathrm{a2}}^4) - W \frac{L}{N_{\mathrm{t}}} h_3 (T_{\mathrm{t}}^{(k)}(t) - T_{\mathrm{am}}) \end{aligned} \tag{3.27}$$

where W, H, and L denote the width, the height, and the length of molten steel in the tundish, respectively. $T_{\mathrm{t}}^{(k)}$ is the tundish temperature in the kth bath. $S_{\mathrm{t}}^{(k)}$ denotes the contact area between the molten steel and

the tundish for the kth bath, U_{t} the overall heat transfer coefficient of the tundish, ε_{t} the emissivity of the molten steel, T_{a2} the air temperature in the tundish, and h_3 the heat transfer coefficient between the molten steel and the air.

3.3.3 Parameter Fitting

The first-principle model contains 13 parameters to be identified: α_0, γ, η, h_1, h_2, h_3, U_{b0}, U_{t}, T_{a1}, T_{a2}, $\varepsilon_{\mathrm{sl}}$, ε_{t} and β. The first ten parameters were estimated through the least squares algorithm using real process data and $\varepsilon_{\mathrm{sl}}$, ε_{t} and β were given in advance from the experience of engineers. The dimensions of the ladle and the tundish, and the physical properties of steel were also given in advance. N_{t} was set equal to three based on the CFD study by Odenthal et al. [4]. The input variables of the physical model are the number of ladle usage, the weight and the initial temperature of the molten steel in the ladle, the transportation time, the casting flow rate, and the initial temperature of each bath composing the tundish. The output variable is the temperature of the last bath, TD temp. A total of 1270 samples were used for parameter estimation.

3.4 Prediction and Control of Molten Steel Temperature in Tundish

In this section, three types of gray-box models, i.e., the parallel, the serial, and the combined gray-box models, are constructed to predict and control TD temp.

3.4.1 Parallel Gray-box Model of the Process

In the parallel gray-box model, a statistical model to compensate the prediction error of the first-principle model was developed by using 53 process

variables, measured at the processes from the converter to the tundish, including the variables used in the first-principle model. The statistical model was developed by RF. In this RF model, the number of trees and the split features was set at 1000 and 17, respectively.

3.4.2 Serial Gray-box Model of the Process

The prediction performance of the first-principle model might be improved by taking account of the dependence of the parameters on the process conditions. That is, a serial gray-box model might be effective in reducing the prediction errors. However, it is not clear which parameter should be updated according to changes in process conditions. Each of all the parameters was regarded as a candidate variable to be expressed as a function of measured process variables, and the most influential parameter was selected. Thus, the following index was calculated for each parameter after the development of the first-principle model.

$$\mathrm{MAE}_i = \frac{1}{N_\mathrm{d}} \sum_{n=1}^{N_\mathrm{d}} |y_n - f_\mathrm{fp}(X_{\mathrm{fp},n}, \boldsymbol{\theta}_i^c, \theta_{i,n})|, \ (i = 1, 2, \cdots, N_\mathrm{p}) \tag{3.28}$$

where $\theta_{i,n}$ is the optimal value of $\theta_{i,n}$ which is derived by Eq. (3.8). Mean absolute errors (MAE), after optimization of parameters $\boldsymbol{\theta}$ one by one, is shown in Table 3.1. Optimization of U_b achieved the smallest MAE therefore it was selected as the parameter to be updated for each sample. Since U_b0 is supposed to be a constant value, Eq. (3.22) was discarded and U_b was regarded as a parameter instead of U_b0 and γ. **Figure 3.2** shows relation of the optimized values of U_b to the training errors e of the first-principle model. The positive errors e shows that the heat loss calculated by the first-principle model is larger than the actual heat loss in the plant while the negative errors e shows that the heat loss calculated by the first-principle model is smaller

than the actual heat loss in the plant. To precisely model the heat loss, the value of U_b is decreased for positive errors e and increased for negative errors e. It is clear that the errors can be reduced drastically when U_b is estimated properly. Thus, U_b was selected as the variable to be updated as a function of 53 process variables. The optimal value of U_b was calculated for each modeling sample, and a statistical model was developed by RF to estimate U_b. In this RF model, the number of trees and the split features was set at the same values as that of the RF model in the parallel gray-box model.

3.4.3 Combined Gray-box Model of the Process

The prediction performance of the first-principle model, the parallel gray-box model, the serial gray-box model, and the combined gray-box model, was compared by applying them to real process data. For comparison, a statistical model developed by using only RF was also constructed. The total number of samples was 1588; 1270 samples (80%) were used for modeling and the remaining 318 samples (20%) were used for validation. The prediction results are shown in **Figure 3.3**. The plot at the bottom right of Figure 3.3 shows prediction results of TD temp, through the combined gray-box model, for 30 validation samples. The prediction performance was evaluated on the basis of the root-mean-square error (RMSE) and the correlation coefficient (r) between scaled values of the reference TD temp and the predicted TD temp. RMSE of the first-principle model, the RF model, the parallel gray-box model, the serial gray-box model, and the combined gray-box model is 2.73, 2.08, 1.83, 1.81 and 1.74, respectively. The performance of the proposed combined gray-box model is superior to the other models. The RMSE of the combined gray-box model is 36%, 16%, 5% and 4% smaller than that of the first-principle model, the RF model, the parallel gray-box model, and the serial gray-box model, respectively. The results show that the serial gray-box model performs better than the first-principle model, the RF model, and the parallel gray-box model, however there are still a significant amount of

Table 3.1: Mean absolute error MAE of the first-principle model derived by optimizing each of 10 parameters

Parameter	MAE
α_0	1.28
γ	1.50
η	1.94
h_1	2.29
h_2	3.84
h_3	3.71
U_b	3.30×10^{-5}
U_t	1.51
T_a1	1.85
T_a2	0.03

prediction errors. Such prediction errors cannot be compensated merely by updating the parameters because the structure of the first-principle model also cause some limitations. The combined gray-box model uses an additional statistical model to estimate/compensate such prediction errors and therefore outperforms all the models including the serial gray-box model.

3.4.4 Control of Molten Steel Temperature in Tundish

Since the combined gray-box model was able to predict TD temp accurately, the next step was to adjust RH temp in order to realize precise control of TD temp. Thus, a control model was constructed on the basis of the combined gray-box model. The control model derives RH temp from the target TD temp $T_\mathrm{TD}^\mathrm{set}$ and other information available.

The control model is actually an inverse form of the developed model. In the inverse model the task is to provide suitable value of the manipulated variable RH temp in order to get the target TD temp. For this purpose RH temp is optimized with the bisection method. Three different values of RH temp, i.e., RH-max, RH-mid and RH-min, are determined. Initially, TD

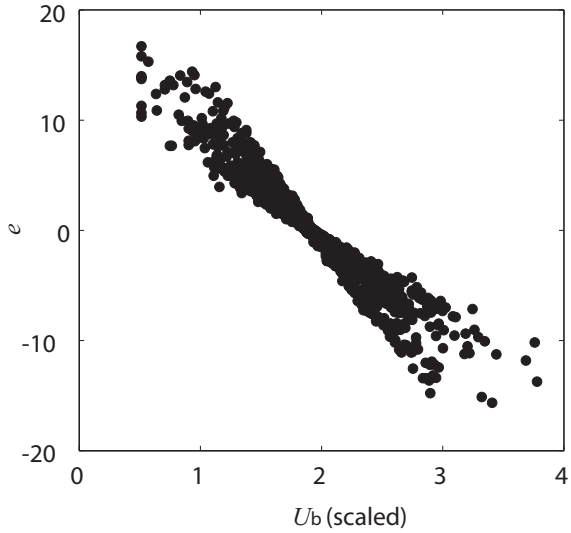


Figure 3.2: Training errors of the first-principle model and the corresponding optimized values of U_b

temp is predicted through the combined gray-box model using RH-mid. Then the value of RH temp is iteratively updated using the bisections method and new TD temp is predicted until it becomes equal to the target TD temp. **Figure 3.4** shows the results of RH temp derivation by using the first-principle model, the RF model, the parallel gray-box model, the serial gray-box model, and the combined gray-box model. The plot at the bottom right of Figure 3.4 shows derivation results of RH temp, through the combined gray-box model, for 30 validation samples.

The performance was evaluated on the basis of the root-mean-square error (RMSE) and the correlation coefficient (r) between scaled values of the reference RH temp and the derived RH temp. RMSE of the first-principle model, the RF model, the parallel gray-box model, the serial gray-box model, and the combined gray-box model is 2.90, 1.65, 1.95, 1.80 and 1.64, respectively. The performance of the proposed combined gray-box model is superior to the other models. The combined gray-box model achieved the highest prediction accuracy and its RMSE is 43%, 1%, 16% and 10% smaller than those

of the first-principle model, the RF model, the parallel gray-box model, and the serial gray-box model, respectively. In the control model, the RF model outperforms the other models except the combined gray-box model. The reason is that RF model does not use the inverse approach followed by the other models; it is simply trained on the TD temp, $T_{\mathrm{TD}}^{\mathrm{set}}$ and other measured variables, to predict RH temp. Although RF can build a nonlinear process model, its direct application is not always the best approach because of its weaker interpretability than the gray-box models.

3.5 Conclusions

In the steel making process, the molten steel temperature in the tundish (TD temp), which is one of the key factors to realize stable operation, is controlled by the molten steel temperature in the secondary refining process (RH temp). In this chapter, a new type of gray-box model is proposed and was applied to predict TD temp in a steel making plant. The feature of the proposed gray-box model is to predict the parameters in the first-principle model by statistical models. With such a model structure, it becomes possible to adjust a parameter as a function of measured variables. The developed gray-box model, called the serial gray-box model, was further combined with a parallel gray-box model in which the prediction errors of the first-principle model was compensated by the statistical model. The combination of the serial and the parallel gray-box model resulted in a combined gray-box model.

To use the proposed gray-box model in controlling of TD temp, it was inversely used to derive temperature of the Ruhrstahl-Heraeus degassing process (RH temp) at the end of its operation from the desired tundish TD temp. The results in TD temp prediction as well as RH temp derivation show the advantage of the proposed gray-box models over the first-principle model, the statistical model and the parallel gray-box model.

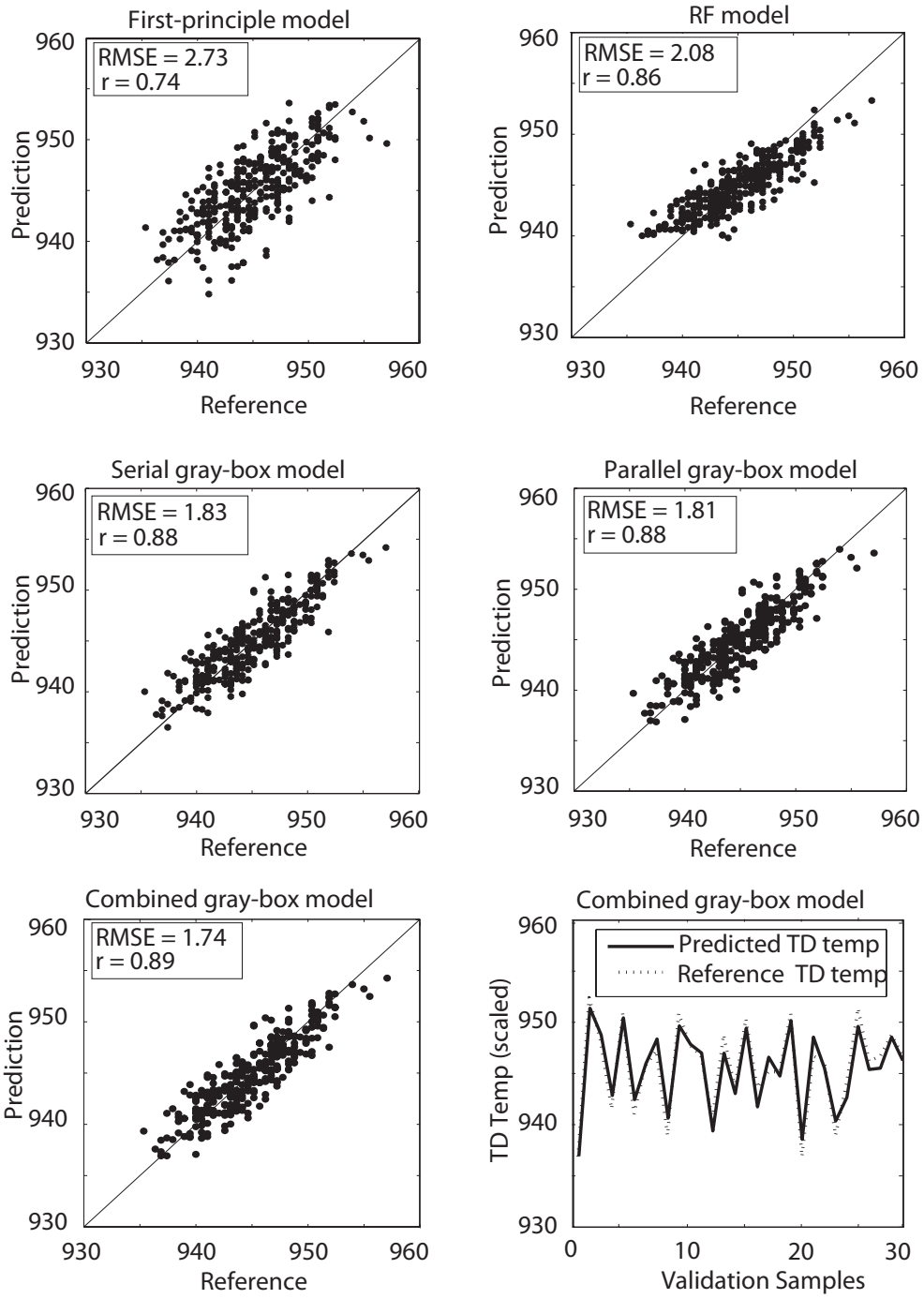


Figure 3.3: Prediction of TD temp through first-principle model, RF model, parallel gray-box model, serial gray-box model, combined gray-box model, and predicted and reference TD temp

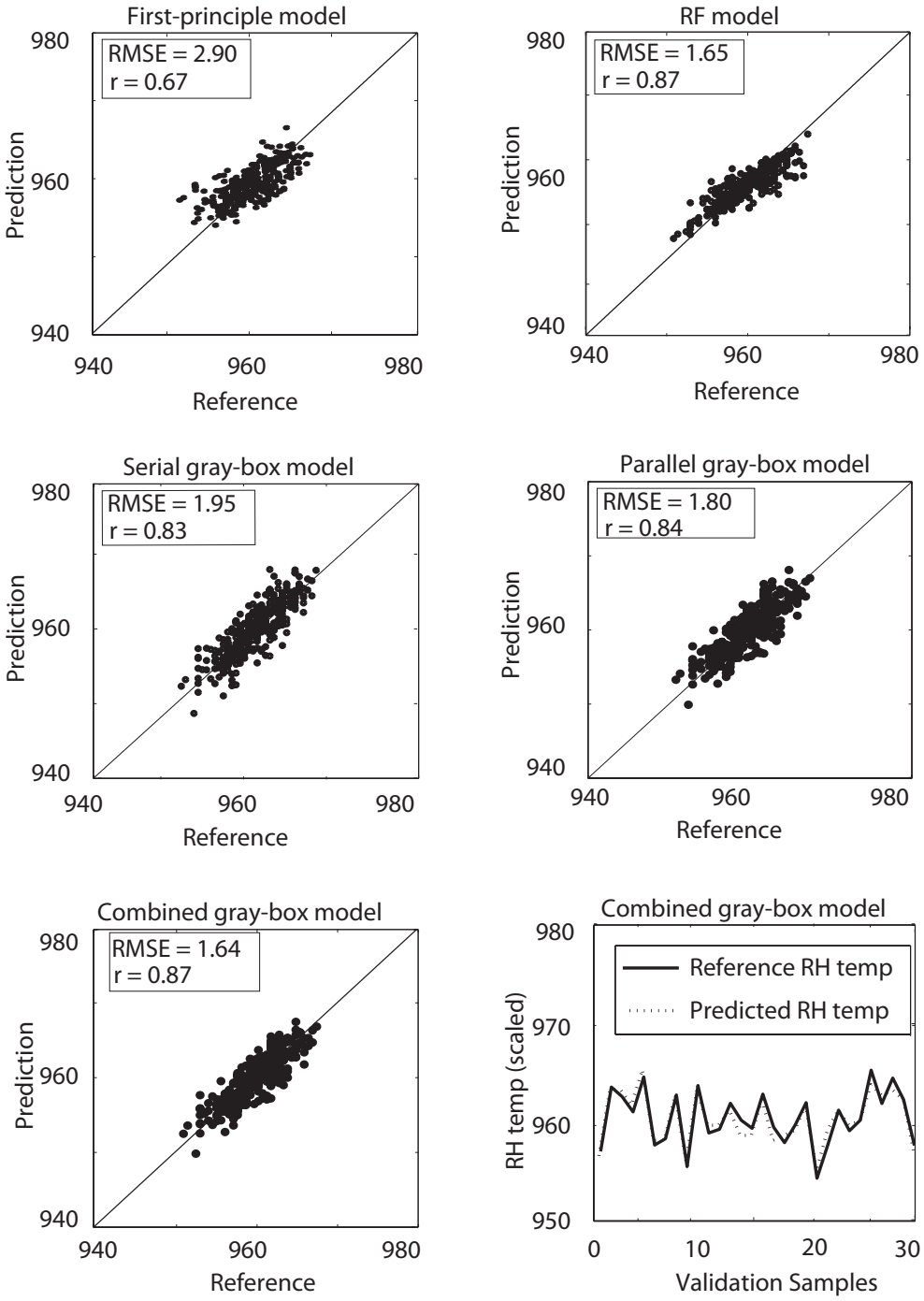


Figure 3.4: Derivation of RH temp through first-principle model, RF model, parallel gray-box model, serial gray-box model, combined gray-box model, and derived and reference RH temp

Bibliography

[1] T. A. Johansen, B. A. Foss: Representing and learning unmodelled dynamics with neural network memories, *American Control Conference*, (1992), 3037-3043.

[2] D. C. Psichogios: L. H. Ungar: A hybrid neural network first principles approach to process modeling, *AIChE Journal*, 38 (10) (1992), 1499-1511.

[3] T. Okura, I. Ahmad, M. Kano, S. Hasebe, H. Kitada, N. Murata: High-performance prediction of molten steel temperature in tundish through gray-box model, *ISIJ International*, 53 (1) (2013), 76-80.

[4] H. J. Odenthal, M. Javurek, M. Kirschen, N. Vogl: CFD benchmark for a single strand tundish, *Steel Research International*, 81 (7) (2010), 529-541.

Chapter 4

Modeling Uncertainty in Steel Making

4.1 Introduction

Stable operation of a continuous casting process requires precise control of molten steel temperature in a tundish (TD temp), which is a container used to feed molten steel into an ingot mold. Since TD temp is implicitly controlled by adjusting molten steel temperature in the preceding secondary refining process (RH temp), a model relating TD temp with RH temp is required. This chapter proposes a procedure to predict the probability distribution of TD temp by integrating a gray-box model and a bootstrap filter to cope with uncertainties of the process. The derived probability distribution is used not only to predict TD temp but also to evaluate the reliability of prediction.

Although deterministic models are dominant in the literature [1]−[6], a stochastic model is preferable to cope with process uncertainties, such as those in temperature measurements, composition and weight of added alloys, the extent of oxidation reactions for removal of impurities, and degradation of ladles. Since these uncertainties are unavoidable in the steel making process, they should be embedded in the model to evaluate their influence on TD temp

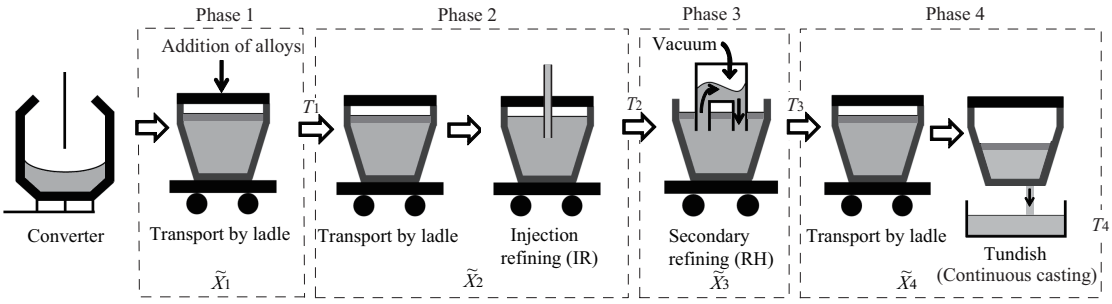


Figure 4.1: Steel making process divided into four phases according to observations of molten steel temperature

prediction. In this research, we treated the molten steel temperature in phase i, T_i $(i = 1, 2, 3, 4)$, in **Figure 4.1** as a stochastic variable, and we estimated TD temp (T_4) in the form of a probability distribution by integrating a gray-box model with a bootstrap filter. In the proposed method, the gray-box model is used as a deterministic function for the bootstrap filter, in which Bayes' theorem is effectively used to adjust the probability distribution of temperature. As a result, the risk caused by process uncertainties can be evaluated, and feasible operation can be suggested on the basis of the probability distribution of the TD temp.

This chapter is organized as follows. In section 4.2, models used in the research are described. Then, the procedure of predicting probability distribution of TD temp is explained in section 4.3. The prediction results are shown in section 4.4, followed by the conclusions in section 4.5.

4.2 Mathematical Modeling

In this section, the procedure of building a gray-box model of the steel making process is explained.

4.2.1 Process and Data

In steel making, a converter is used to make carbon-rich molten pig iron into steel. Then, as shown in Figure 4.1, the molten steel in the converter is discharged to a ladle, and ferroalloys are added to the molten steel in order to adjust the composition. The ladle is a vessel used to transport and pour out molten steel. The injection refining (IR) is a post treatment process of molten steel where some material is injected in powder form with a high pressure for deoxidation, desulphurization and inclusion modification. In a secondary refining process, also called the Ruhrstahl-Heraeus (RH) process, operations such as desulfurization, degassing of oxygen, nitrogen, hydrogen, etc., removal of inclusions, and decarburization are performed in the gas stirred ladle with electric arc heating in the lid of the furnace. After the secondary refining, the molten steel is transported to a continuous casting process and discharged to a tundish.

In the target plant, the molten steel temperature is measured at the end of the four phases shown in Figure 4.1: temperature in the ladle just after the addition of ferroalloys into the ladle (T_1), temperature in the ladle at the beginning and at the end of RH process (T_2, T_3), and temperature in the tundish after the discharge of the molten steel from the ladle to the tundish (T_4). Here, T_4 is TD temp, which is the target temperature of prediction. In the following sections, observed temperature is denoted by $\tilde{T}_i$ and real (unobserved) temperature is denoted by T_i. Other variables measured in phase i are denoted by X_i and their observations by $\tilde{X}_i$. The total number of samples was 1588; 1270 samples (80%) were used for modeling and the remaining 318 samples (20%) were used for validation.

4.2.2 Model between Phases

In the proposed prediction method, the molten steel temperature T_i ($i = 1, 2, 3, 4$) is calculated from T_{i-1} and X_i through the following models:

$$T_1 = f_1(X_1) + \varepsilon_1 \tag{4.1}$$

$$T_i = f_i(T_1, \cdots, T_{i-1}, X_1, \cdots, X_i) + \varepsilon_i \qquad (i = 2, 3, 4) \tag{4.2}$$

where f_i is a deterministic function, and ε_i denotes a system noise, which is assumed to be a white noise following the Gaussian distribution with mean 0 and variance $\sigma^2_{\varepsilon,i}$:

$$P(\varepsilon_i) = N(0, \sigma^2_{\varepsilon,i}) \tag{4.3}$$

Any model can be used as a deterministic function; a gray-box model is used in phase 4 while statistical models are used in the other phases 1, 2, and 3. The gray-box model is used only in phase 4, because this phase includes TD temp T_4 and is the most complicated and important. The gray-box model to predict T_4 is explained in detail in section 4.2.4. The statistical models to predict T_i ($i = 1, 2, 3$) are developed by using random forests (RF).

4.2.3 Derivation of Variances

To quantitatively express process uncertainties in phase 1, the probability distribution of $\tilde{X}_1$ is assumed to be represented by using artificial samples, which are generated from the uniform distribution of $[0.95\tilde{X}_1, 1.05\tilde{X}_1]$. The uniform distribution is used because there is no information about the probability distribution. Ideally, process uncertainties should be quantified by identifying all sources and by determining their contribution from the information provided by the quality control section [7]. However, such quantification of uncertainties is not always possible due to the limitations of available information. In the present work, it is assumed that all input variables of f_1, i.e., X_1, have the same degree of uncertainties and such uncertainties include

the system noise, i.e., $\varepsilon_1 = 0$.

The initial probability distribution of T_1 is derived through Eq. (4.1) with artificial samples of X_1. The observed temperature $\tilde{T}_i$ is regarded as a realization of T_i and described by the following observation model:

$$\tilde{T}_i = T_i + \xi_i \tag{4.4}$$

where ξ_i is an observation error following the Gaussian distribution with mean 0 and variance $\sigma^2_{\xi,i}$.

The variances $\sigma^2_{\varepsilon,i}$ and $\sigma^2_{\xi,i}$ are estimated from the observed temperature $\tilde{T}_i$. First, given $[\tilde{X}_1, \cdots, \tilde{X}_i]$, J nearest neighbors of $\tilde{X}_i$ are selected from the modeling dataset by using k-d tree [8], and the selected neighbors are expressed by $\{\tilde{T}_1(j), \cdots, \tilde{T}_i(j),\ \tilde{X}_1(j), \cdots, \tilde{X}_i(j)\}$ $(j = 1, \cdots, J)$. Then, the prediction error $e_i(j)$ is calculated for all the neighbor samples.

$$e_1(j) = \tilde{T}_i(j) - f_1(\tilde{X}_1(j)) \tag{4.5}$$

$$e_i(j) = \tilde{T}_i(j) - f_i(\tilde{T}_1(j), \cdots, \tilde{T}_{i-1}(j), \tilde{X}_1(j), \cdots, \tilde{X}_i(j)) \quad (i = 2, 3, 4) \tag{4.6}$$

This prediction error is the sum of the system noise and the observation error. In this research, it is assumed that the variances of them are the same and given by

$$\sigma^2_{e,i} = \frac{1}{J} \sum_{j=1}^{J} e_i^2(j) \tag{4.7}$$

$$\sigma^2_{\varepsilon,i} = \sigma^2_{\xi,i} = \frac{1}{2} \sigma^2_{e,i} \tag{4.8}$$

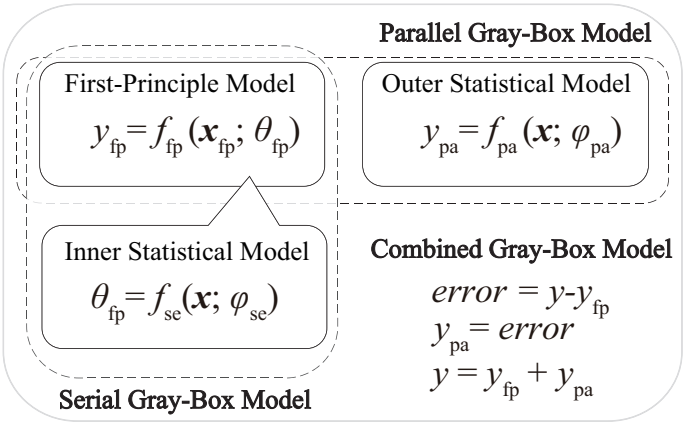


Figure 4.2: Framework of gray-box modeling

4.2.4 Combined Gray-box Model

In general, a gray-box model is a combination of a first-principle model (or a physical model) and a statistical model. We adopted the combined gray-box model recently proposed by Ahmad et., al., (2013) because the prediction accuracy is higher than that of other models.

The generalized framework of gray-box modeling is shown in **Figure 4.2**. When an (outer) statistical model is directly used to compensate the prediction error of a first-principle model, such a gray-box model is called a parallel gray-box model. On the other hand, it is called a serial gray-box model when an (inner) statistical model is used to update parameters of a first-principle model. A combined gray-box model is a combination of a serial gray-box model and a parallel gray-box model; it consists of a first principle model, an inner statistical model, and an outer statistical model. To develop such a combined gray-box model, a serial gray-box model is developed first, then its prediction error is compensated with an outer statistical model. The overall modeling procedure is as follows:

1. Build a first-principle model

$$y_{\mathrm{fp}} = f_{\mathrm{fp}}(\boldsymbol{x}_{\mathrm{fp}}, \boldsymbol{\theta}) \tag{4.9}$$

where y_{fp} is an estimate of an output variable y, $\boldsymbol{x}_{\mathrm{fp}}$ and $\boldsymbol{\theta}$ are input variables and parameters of the first-principle model, respectively.

2. Estimate $\boldsymbol{\theta}$ that minimizes the sum of squared errors.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^{N_{\mathrm{s}}} \left(y(i) - f_{\mathrm{fp}}(\boldsymbol{x}_{\mathrm{fp}}(i), \theta)\right)^2 \tag{4.10}$$

$$\theta_k^{\mathrm{L}} \leq \theta_k \leq \theta_k^{\mathrm{U}} \qquad (k = 1, 2, \cdots, N_{\mathrm{p}}) \tag{4.11}$$

where $\boldsymbol{x}_{\mathrm{fp}}(i)$ and $y(i)$ are the ith measurements of input and output variables, respectively. N_{s} is the number of samples for modeling, and N_{p} is the number of parameters. θ_k^{L} and θ_k^{U} are lower and upper bounds of the kth parameter.

3. Select a parameter θ_k $(k = 1, 2, \cdots, N_{\mathrm{p}})$ and optimize it for each modeling sample under the condition that the parameters except for θ_k are fixed at the optimal values obtained at step 2.

$$\hat{\theta}_k(i) = \underset{\theta_k(i)}{\operatorname{argmin}} \left(y(i) - f_{\mathrm{fp}}(\boldsymbol{x}_{\mathrm{fp}}(i), \hat{\boldsymbol{\theta}}_{(k)}, \theta_k(i))\right)^2 \qquad (i = 1, 2, \cdots, N_{\mathrm{s}}) \tag{4.12}$$

$$\theta_k^{\mathrm{L}} \leq \theta_k \leq \theta_k^{\mathrm{U}} \tag{4.13}$$

where $\hat{\theta}_k(i)$ is the optimal value of the parameter θ_k for the ith modeling

sample, and $\hat{\boldsymbol{\theta}}_{(k)}$ is the constant vector consisting of the optimal parameters except for the kth parameter.

4. Repeat step 3 for all parameters one at a time, and select the Kth parameter that minimizes the absolute error MAE_k:

$$K = \underset{k}{\operatorname{argmin}}\, MAE_k \qquad (k = 1, 2, \cdots, N_p) \tag{4.14}$$

$$MAE_k = \frac{1}{N_s} \sum_{i=1}^{N_\mathrm{s}} \mid y(i) - f_\mathrm{fp}(\boldsymbol{x}_\mathrm{fp}(i), \hat{\boldsymbol{\theta}}_{(k)}, \hat{\boldsymbol{\theta}_k}(i)) \mid \tag{4.15}$$

5. Build a statistical model $f_{\mathrm{se},K}$ to estimate the optimal Kth parameter $\hat{\theta}_K$ from input variables $\boldsymbol{x}$.

$$\hat{\boldsymbol{\varphi}}_{\mathrm{se},K} = \underset{\boldsymbol{\varphi}_{\mathrm{se},K}}{\operatorname{argmin}} \sum_{i=1}^{N_\mathrm{s}} \left(\hat{\theta}_K(i) - f_{\mathrm{se},K}(\boldsymbol{x}(i), \boldsymbol{\varphi}_{\mathrm{se},K}) \right)^2 \tag{4.16}$$

where $\boldsymbol{\varphi}_{\mathrm{se},K}$ is a parameter vector of the statistical model. The input variables, which are not included in the first principle model, can be used in $\boldsymbol{x}$; in general, the variables in $\boldsymbol{x}_\mathrm{fp}$ are a part of those in $\boldsymbol{x}$.

6. Derive a gray-box model by combining the first-principle model and the statistical model.

$$y_\mathrm{se} = f_\mathrm{fp}(\boldsymbol{x}_\mathrm{fp}, \hat{\boldsymbol{\theta}}_{(K)}, \hat{\boldsymbol{\theta}}_K) \tag{4.17}$$

$$\hat{\boldsymbol{\theta}}_K = f_{\mathrm{se},K}(\boldsymbol{x}, \hat{\boldsymbol{\varphi}}_{\mathrm{se},K}) \tag{4.18}$$

$$e_\mathrm{se} = y - y_\mathrm{se} \tag{4.19}$$

where y_se and e_se are the output and the prediction error of the serial gray-box model.

1. Build a statistical model f_{pa} to predict the output error e_{se} from $\boldsymbol{x}$.

$$\hat{\boldsymbol{\varphi}}_{\mathrm{pa}} = \underset{\boldsymbol{\varphi}_{\mathrm{pa}}}{\operatorname{argmin}} \sum_{i=1}^{N_{\mathrm{s}}} \left(e_{\mathrm{se}}(i) - f_{\mathrm{pa}}(\boldsymbol{x}(i), \boldsymbol{\varphi}_{\mathrm{pa}})\right)^2 \tag{4.20}$$

   where $\boldsymbol{\varphi}_{\mathrm{pa}}$ is parameters of the statistical model.

2. Build a combined gray-box model by combining the serial gray-box model and the statistical model.

$$y_{\mathrm{com}} = f_{\mathrm{fp}}(\boldsymbol{x}_{\mathrm{fp}}, \hat{\boldsymbol{\theta}}_{(K)}, \hat{\boldsymbol{\theta}}_K) + f_{\mathrm{pa}}(\boldsymbol{x}, \hat{\boldsymbol{\varphi}}_{\mathrm{pa}}) \tag{4.21}$$

   where y_{com} is the prediction of the target output variable by using the combined gray-box model.

4.2.5 First-Principle Model

After the secondary refining is finished, the molten steel is transported to the continuous casting process with a ladle. It is then discharged from the ladle to the tundish. To predict the temperature in the tundish, TD temp, with the gray-box model, the first-principle model plays an important role. The first-principle model consists of two sub-models. The first sub-model describes the phenomena during the transportation from the secondary refining process to the continuous casting process. The second sub-model describes the phenomena during the discharging of molten steel from the ladle to the tundish. The first-principle model adopted here is basically the same as the model developed by Okura et., al., (2013).

The first-principle model contains 13 parameters, which are summarized in **Table 4.1**. Three parameters $\varepsilon_{\mathrm{sl}}$, ε_{t}, and β, are determined in advance from engineers' experience, and the remaining ten parameters are estimated through the least squares method using real process data. The input variables of the first-principle model are the number of ladle usage, the weight and the

Table 4.1: Parameters in the first-principle model

α, η:	parameters which are related to the temperature difference between the top and bottom of a ladle
h_1, h_2, h_3:	heat transfer coefficients between the molten steel and the slag, the slag and the air in the ladle, and the molten steel and the air, respectively
U_{b0}:	initial value of the overall heat transfer coefficient of the ladle bottom
U_{t}:	overall heat transfer coefficient of the tundish
T_{a1}, T_{a2}:	air temperatures in the ladle and in the tundish
$\varepsilon_{\mathrm{sl}}$, ε_{m}:	emissivity of the slag and the molten steel
β, γ:	parameters which are related to the overall heat transfer coefficients of the ladle bottom and ladle wall

initial temperature of the molten steel in the ladle, the transportation time, the casting flow rate, and the initial temperature of each bath composing the tundish. The output variable is TD temp, T_4, which is the temperature of the last bath in the tundish model.

4.2.6 Random Forests

In this research, the molten steel temperature T_i $(i = 1, 2, 3)$ is predicted through Eqs. (4.1) and (4.2) with random forests (RF). X_1 consists of 22 variables such as temperature, number of ladle use, and weight of alloys added in phase 1, X_2 consists of 16 variables such as transportation time, weight of alloys added in IR and amount of oxygen used in IR in phase 2, X_3 consists of four variables such as temperature and circulation time in RH degassing in phase 3, and X_4 consists of 11 variables such as transportation time, casting time and wight of molten steel. The number of samples $N = 1270$, and bagging generates $M = 1000$ new training sets for each model. The numbers of trees and the split features are set at the same values for all

phases, and they are 1000 and 17, respectively.

4.3 Probability Distribution of Temperature in Tundish

In this section, a procedure of predicting the probability distribution of TD temp, is proposed, and a method of evaluating the reliability of the prediction is explained. To assess the usefulness of the combined gray-box model in phase 4, a statistical model is developed with RF and compared.

4.3.1 Prediction Procedure of Probability Distribution of Temperature in Tundish

To quantify the influence of uncertainties in the input variables on the variation of output variables, we adopted one of particle filters called bootstrap filter [11]. The developed procedure of predicting the probability distribution of temperatures consists of the following steps:

1. Take a set of measured data of the molten steel temperatures $\tilde{T}_i$ $(i = 1, 2, 3)$ and the input variables $\tilde{X}_i$ $(i = 1, 2, 3, 4)$. The planned operating condition is assigned to $\tilde{X}_4$.

2. Calculate the variance of system noise $\sigma^2_{\varepsilon,i}$ $(i = 2, 3)$ and that of observation error $\sigma^2_{\xi,i}$ $(i = 1, 2, 3)$ through the procedure explained in section 4.2.3. In this work, the number of nearest neighbors J is set at 30.

3. Generate N_a artificial samples of $\tilde{X}_1$, $\tilde{X}_1^{(n)}$ $(n = 1, 2, \cdots, N_\mathrm{a})$, from the uniform distribution of $[0.95\tilde{X}_1, 1.05\tilde{X}_1]$. In this work, $N_\mathrm{a} = 200$.

4. Set $i = 1$.

5. Predict temperature $T_i^{(n)}$ $(n = 1, 2, \cdots, N_\mathrm{a})$.

$$T_1^{(n)} = f_1(\tilde{X}_1^{(n)}) \tag{4.22}$$

$$T_i^{(n)} = f_i(T_1^{(n)}, \cdots, T_{i-1}^{(n)}, \tilde{X}_1^{(n)}, \cdots, \tilde{X}_i^{(n)}) + \varepsilon_i^{(n)} \quad (i = 2, 3, 4) \tag{4.23}$$

   where $\varepsilon_i^{(n)}$ is drawn from $P(\varepsilon_i)$ in Eq. (4.3).

6. Calculate the likelihood of $T_i^{(n)}$ through

$$W_i^{(n)} = p(T_i^{(n)}; \tilde{T}_i, \sigma_{\xi,i}^2) = \frac{1}{\sqrt{2\pi\sigma_{\xi,i}^2}} \exp\left(-\frac{(T_i^{(n)} - \tilde{T}_i^2)}{2\sigma_{\xi,i}^2}\right) \tag{4.24}$$

   where $p(z; \mu, \sigma^2)$ is the likelihood of z for the Gaussian distribution with mean μ and variance σ^2.

7. Resample N_a samples of temperature $T_i^{+(n)}$ from $T_i^{(n)}$ so that the probability of $T_i^{+(n)}$ is proportional to $W_i^{(n)}$ $(n = 1, 2, \cdots, N_\mathrm{a})$. The resampling reshapes the prior distribution from the information of $\tilde{T}_i$ and derives the posterior distribution.

8. Unless $i = 4$, set $i = i + 1$ and go to step 5. Otherwise, go to step 9.

9. Calculate the expectation value of T_4 and its variance by using the samples obtained at step 7.

An example of the prediction of the probability distribution of temperatures is shown in **Figure 4.3**, which demonstrates that the variance of T_i can be reduced by executing step 7.

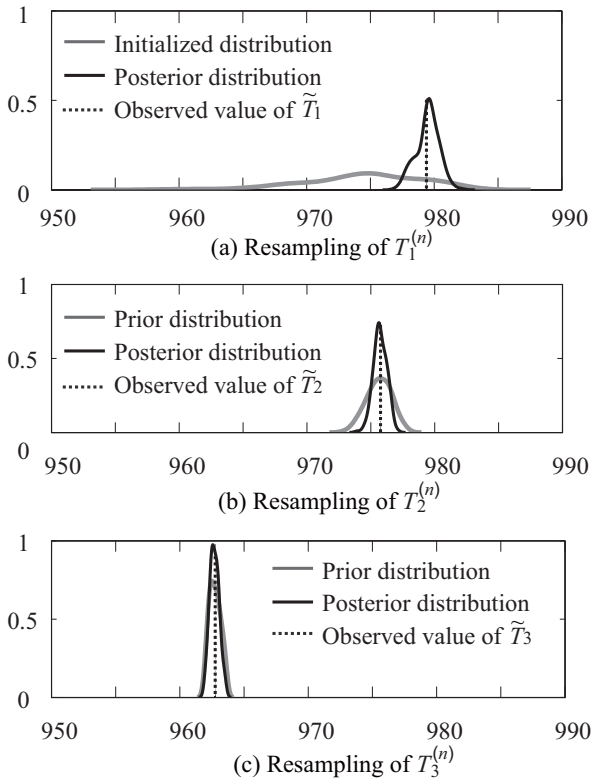


Figure 4.3: Prior and posterior distributions and observed values of molten steel temperatures in phases 1, 2, and 3

4.3.2 Reliability Check of Predicted Probability Distribution

A major objective of predicting TD temp, is to use the predicted value for process control. Since the control performance depends on the accuracy of the prediction, it is important to evaluate the reliability of prediction. The probability distribution of T_4 can be effectively used for this purpose.

To determine control limits on T_4, the nearest neighbor sample of the target sample $[\tilde{T}_1, \tilde{T}_2, \tilde{T}_3, \tilde{X}_1, \tilde{X}_2, \tilde{X}_3, \tilde{X}_4]$, for which the prediction is required, is selected from the modeling samples by using k-d tree. Based on the observed molten steel temperature of the nearest neighbor sample, $\tilde{T}_4^{\mathrm{N}}$, upper and

lower control limits are set at $1.01\tilde{T}_4^{\mathrm{N}}$ and $0.99\tilde{T}_4^{\mathrm{N}}$, respectively. Then, the probability P_{lim} that the prediction is within the control limits are calculated:

$$P_{\mathrm{lim}} = \int_{0.99\tilde{T}_4^{\mathrm{N}}}^{1.01\tilde{T}_4^{\mathrm{N}}} p(T_4)dT_4 \tag{4.25}$$

where $p(T_4)$ is the probability density function of T_4. To evaluate the reliability of prediction by using P_{lim}, the following rule can be used.

$$\begin{cases} P_{\mathrm{lim}} > \Theta : \text{prediction is reliable} \\ P_{\mathrm{lim}} \leq \Theta : \text{prediction is not reliable} \end{cases} \tag{4.26}$$

where Θ is a threshold determined by engineers based on their experience. When $P_{\mathrm{lim}} \leq \Theta$, operators are expected to change the operation conditions to avoid risk of violating steel product quality requirements.

4.3.3 Validation of the Control Limits

To check the reliability of prediction, control limits are determined by using the molten steel temperature of the nearest neighbor sample $\tilde{T}_4^{\mathrm{N}}$ just after a new sample (query) is obtained. This just-in-time (JIT) approach may not function well when there is no sample close to the query, because the nearest neighbor sample is not a good approximation of the query sample. To validate the accuracy of the control limits derived by the JIT approach, we used the observed molten steel temperature $\tilde{T}_4$ of the validation samples to construct the reference control limits: $0.99\tilde{T}_4$ and $1.01\tilde{T}_4$. By using these reference control limits, the reliability of prediction is also evaluated.

4.4 Results

This section provides the results of applying the proposed method to real plant data.

Table 4.2: Mean absolute error MAE_k of the first-principle model derived by optimizing each of 10 parameters

Parameter	MAE_k
α_0	1.28
γ	1.50
η	1.94
h_1	2.29
h_2	3.84
h_3	3.71
U_b	3.30×10^{-5}
U_t	1.51
T_a1	1.85
T_a2	0.03

4.4.1 Combined Gray-box Model

The first-principle model has 13 parameters. To build a serial gray-box model, the most effective parameter is selected by Eq. (4.12) and expressed as a function of measured process variables with RF in order to improve the prediction performance. Since three parameters are predetermined from the experience of engineers, the remaining 10 parameters listed in **Table 4.2** are candidates to be estimated by a statistical model in a serial gray-box model. The mean absolute error MAE_k, after optimizing the parameters for training data, is summarized in Table 4.2. The smallest MAE_k was obtained by optimizing U_b, therefore U_b was selected as the most effective parameter. The optimal value of U_b was calculated for each modeling sample, and a statistical model was developed by RF to estimate U_b from 53 process variables.

To compensate the remaining prediction error, another RF model was constructed and combined with the serial gray-box model. As a result, the combined gray-box model was developed. In two RF models used in the combined gray-box model, the numbers of trees and the split features were set at the same values of 1000 and 17.

Table 4.3: Comparison of the combined gray-box model and the RF model in the prediction performance

Criteria	Model type	Phase 2	Phase 3	Phase 4
RMSE	Gray-box	-	-	2.20
	RF	4.50	1.91	2.36
r	Gray-box	-	-	0.86
	RF	0.85	0.87	0.82

4.4.2 Prediction Results

In addition to the combined gray-box model, we developed a statistical model to predict TD temp with RF and compared them. The prediction performance of the proposed combined gray-box model and the RF model was validated with real process data. In the RF model, 53 process variables were used as input variables, and the number of trees and the split features was 1000 and 17, respectively.

Single-point predictions of the scaled molten steel temperatures in phases 2, 3, and 4 were calculated as the expectation values from the corresponding probability distributions. The prediction performance was evaluated with the root-mean-square error (RMSE) and the correlation coefficient (r) between the measured (reference) values and the predicted expectation values. The prediction results shown in **Table 4.3** and **Figure 4.4(a, b)** confirm that the combined gray-box model outperformed the RF model in predicting TD temp, T_4.

The reliability of the predicted distribution of T_4 by gray-box model P_{lim}, for 318 validation samples are shown in Figure 4.4 (c). In this figure, samples are sorted by the value of P_{lim}. For the gray-box model, 95% of the probability distribution lie within the control limits when the threshold Θ is 0.50. It is confirmed that the probability P_{lim} was very small for some samples; that means the prediction was out of control limits with a high probability for

them. This is valuable information for engineers and operators.

The JIT-based control limits and the reference control limits on T_4 for the first 50 validation samples are shown in Figure 4.4 (d, e). Among 318 validation samples, 15 samples (5%) were out of control when the JIT approach was adopted. On the other hand, no samples were out of control when the observed values $\tilde{T}_4$ were used to determine the control limits (reference control limits). The results demonstrate that the prediction of the proposed combined gray-box model was reliable.

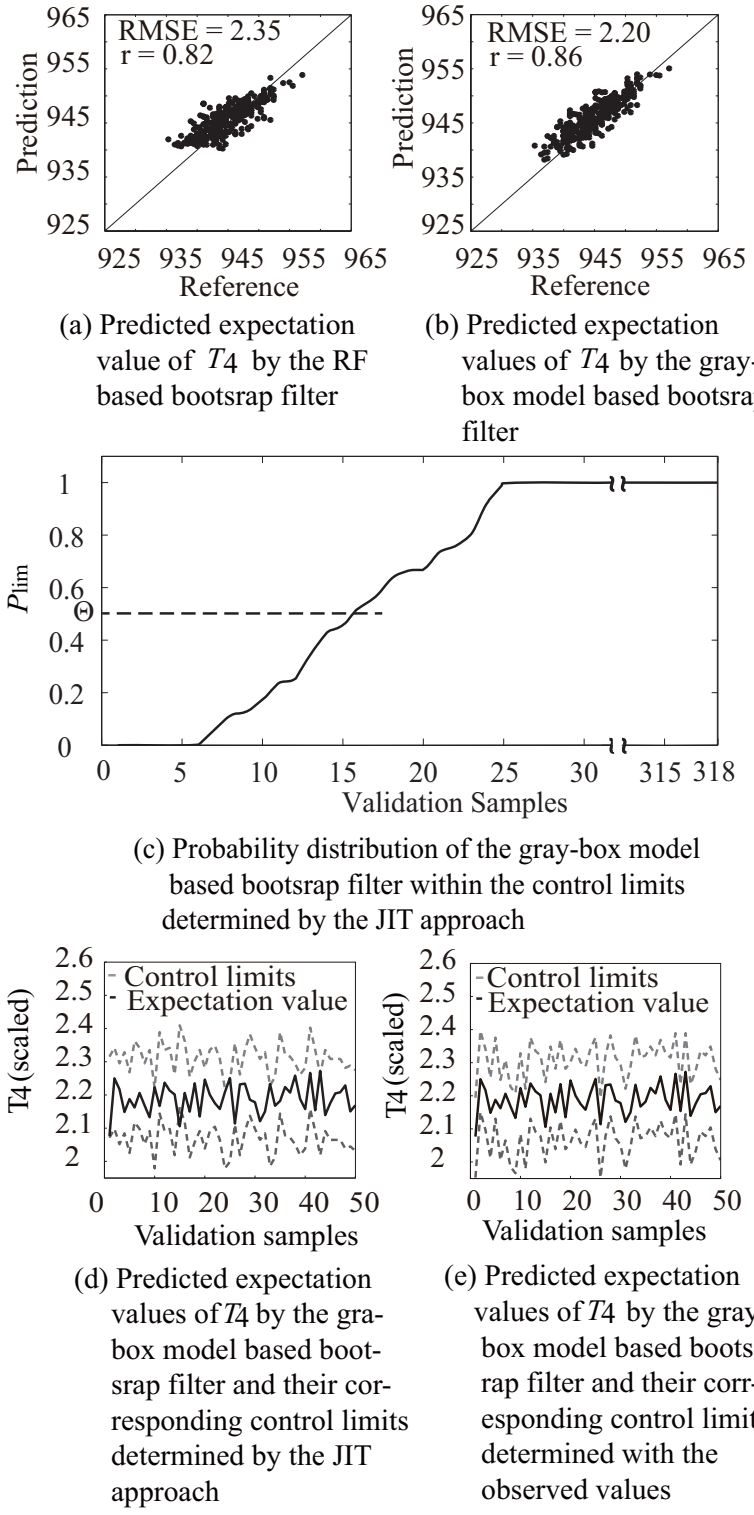
4.5 Conclusions

In the steel making process, the molten steel temperature in the tundish (TD temp) is one of the key factors for realizing stable operation. Various process uncertainties affect TD temp and deteriorate the quality of steel products. In this research, we proposed a procedure for deriving the probability distribution of predicted TD temp. First, the plant was divided into four phases, and a statistical or gray-box model that predicted the molten steel temperature at the end of each phase was constructed. Then, these models were combined with the bootstrap filter to calculate the posterior distribution of the temperature prediction.

The prediction accuracy of the proposed combined gray-box model was compared with that of the statistical model based on RF. The results of applying the two models to real plant data indicated that the gray-box model was better than the RF model.

To minimize the risk caused by process uncertainties and realize optimal operation, the reliability of prediction was evaluated on the basis of its probability distribution. The JIT approach was proposed to determine upper and lower control limits and evaluate the reliability by using nearest neighbor samples. In the proposed approach, the probability that the prediction was within the control limits was calculated, and it was effectively used to avoid

the risk of violating steel product quality requirements. The proposed modeling strategy was validated through its application to real operation data at a steel works. The prediction performance of the combined gray-box model satisfied the requirement for its industrial application.



(a) Predicted expectation value of $T4$ by the RF based bootsrap filter

(b) Predicted expectation values of $T4$ by the gray-box model based bootsrap filter

(c) Probability distribution of the gray-box model based bootsrap filter within the control limits determined by the JIT approach

(d) Predicted expectation values of $T4$ by the gra-box model based boot-srap filter and their cor-responding control limits determined by the JIT approach

(e) Predicted expectation values of $T4$ by the gray-box model based boots-rap filter and their corr-esponding control limits determined with the observed values

Figure 4.4: Prediction of T_4 and the reliability of prediction with the proposed method

Bibliography

[1] A. Zoryk, P. M. Reid: On-line liquid steel temperature control, *Iron and Steelmaker*, 20 (1993), 21-27.

[2] A. G. Belkovskii, Y. L. Kats: Mathematical model of the cooling of steel in a small ladle, *Metallurgist*, 53 (2009), 261-273.

[3] P. R. Austin, J. M. Camplin: J. Herbertson, I. J. Taggart, Mathematical modeling of thermal stratification and drainage of steel ladles, *ISIJ International*, 32 (2) (1992), 196-202.

[4] J. L. Xia, T. Ahokainen: Transient flow and heat transfer in a steelmaking ladle during the holding period, *Metallurgical and Materials Transactions B*, 32 (2001), 733-741.

[5] J. R. S. Zabadal, M. T. M. B. Vilhena, S. Q. B. Leit: Heat transfer process simulation by finite differences for online control of ladle furnaces, *Ironmaking and Steelmaking*, 31 (2004), 227-232.

[6] T. Jormalainen, S. Louhenkilpi: A model for predicting the melt temperature in the ladle and in the tundish as a function of operating parameters during continuous casting, *Steel Research International*, 77 (2006), 472-484.

[7] S. L. R. Ellison, A. Williams, (Eds).; " *Eurachem/CITAC guide: quantifying uncertainty in analytical measurement*, Third edition, (2012).

[8] J. H. Friedman, J. Bentely, R. A. Finkel: An algorithm for finding best matches in logarithmic expected time, *ACM Transactions on Mathematical Software*, 3 (3)(1977), 209-226.

[9] I. Ahmad, M. Kano, S. Hasebe, H. Kitada, N. Murata; Gray-Box Modeling for Prediction and Control of Molten Steel Temperature in Tundish, *Journal of Process Control*, (accepted).

[10] T. Okura, I. Ahmad, M. Kano, S. Hasebe, H. Kitada, N. Murata: High-performance prediction of molten steel temperature in tundish through gray-box model, *ISIJ International*, 53 (2013), 76-80.

[11] N. J. Gordon, D. J. Salmond, A. F. M. Smith: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings*, 140 (1993), 107-113.

Chapter 5

Conclusions

The steel industry faces stiff competition in the global market, and each steel company has to realize stable and efficient operation and produce high quality products satisfying various customer demand. In steel making, tundish is a vessel used for delivering molten steel from a ladle to a mold in the continuous casting process and control of the molten steel temperature in the tundish (TD temp) is one of the key factors to realizing stable operation. If TD temp is too high, breakouts may occur and cause tremendous increase in maintenance cost and productivity loss. When the temperature is too low, clogging in the tundish nozzle occurs, which causes disruptions in the casting process. However, no effective manipulated variable is available after the secondary refining process to control TD temp. To realize the target TD temp, therefore, it is necessary to adjust the molten steel temperature in the secondary refining process (Ruhrstahl-Heraeus degassing process). The molten steel temperature at the end of secondary refining operation is hereafter called RH temp. To control TD temp by manipulating RH temp, a model relating TD temp and RH temp needs to be constructed. In the past, various models such as first-principle models, statistical models, and gray-box models have been proposed. The gray-box model, which integrates a first-principle model and a statistical model, has attracted researchers' attention by its capability:

known linear/nonlinear phenomena can be embedded in the first-principle model, and an unknown relationship among variables can be embedded in the statistical model by extracting such a relationship from the data. In general, gray-box models are more accurate than simplified first-principle models, less complicated than computational fluid dynamics (CFD) models, and more easily interpreted than statistical models. The present work aims to develop a gray-box modeling strategy that predicts and controls molten steel temperature in a tundish (TD temp). This dissertation presents a sequential development of gray-box modeling technique for prediction and control of molten steel temperature in the continuous casting process. The proposed modeling strategy is divided into three phases explained in chapters 2, 3 and 4 of the dissertation.

In chapter 2, a novel gray-box model is proposed to estimate molten steel temperature in a continuous casting process at a steel making plant by combining a first-principle model and a statistical model. The first-principle model was developed on the basis of computational fluid dynamics (CFD) simulations to simplify the model and to improve estimation accuracy. Since the derived first-principle model was not able to estimate the molten steel temperature in the tundish with sufficient accuracy, statistical models were developed to estimate the estimation errors of the first-principle model through partial least squares (PLS) and random forest (RF). As a result of comparing the three models, i.e., the first-principle model, the PLS-based gray-box model, and the RF-based gray-box model, the RF-based gray-box model achieved the best estimation performance. Thus, the molten steel temperature in the tundish can be estimated with accuracy by adding estimates of the first-principle model and those of the statistical RF model. The proposed gray-box model was applied to the real process data and the results demonstrated its advantage over other models.

In chapter 3, a new type of gray-box model is proposed and was applied to predict and control TD temp in a steel making plant. The feature of the

proposed gray-box model is to predict the parameters in the first-principle model by statistical models. With such a model structure, it becomes possible to adjust a parameter as a function of measured variables. The developed gray-box model, called the serial gray-box model, was further combined with a conventional gray-box model in which the prediction error of the first-principle model was compensated by the statistical model. The combination of the serial and the conventional gray-box model resulted in a combined gray-box model. To use the proposed gray-box model in controlling of TD temp, it was inversely used to derive temperature of the Ruhrstahl-Heraeus degassing process (RH temp) at the end of its operation from the desired tundish TD temp. The results in TD temp prediction as well as RH temp derivation show the advantage of the proposed gray-box models over the first-principle model, the statistical model and the conventional gray-box model. The proposed modeling strategy is validated through its application to real operation data at a steel work.

In chapter 4, the uncertainties which cause errors in the models developed in chapters 2 and 3 are emphasized. The issue of uncertainties in equipments characteristics, operating conditions, and raw materials in a steel making process is addressed by adopting a new modeling approach. A gray-box model is combined with a bootstrap filter, in order to cope with process uncertainties. The gray-box model is used as a deterministic transition function for the bootstrap filter. With the new modeling strategy, it becomes possible to predict probability distribution of the TD temp. Furthermore, to minimize the risk caused by process uncertainties and realize the optimal operation, the reliability of prediction is evaluated on the basis of its probability distribution. A just-in-time (JIT) modeling , i.e., k-d tree, is used to determine nearest neighbor samples from the training database, and upper and lower control limits are set for the prediction of the molten steel temperature. Then probability distribution of the prediction within the control limits is calculated and the reliability is evaluated using a threshold set with the help of the

company engineers based on their experience in quality control of the steel products. The reliability check helps to determine the moment when changes in the operation conditions are vital, in order to avoid risk of violating steel product quality requirements. The proposed modeling strategy is validated through its application to real operation data at a steel work.

Acknowledgements

I would like to express my gratitude to Professor Shinji Hasebe and Professor Manabu Kano for being great mentors, for their time and what I learnt from them. Special thanks to Professor Masahiro Ohshima and Professor Motoaki Kawase for their valuable input in preparing this dissertation.
I would like to thank my earlier mentors in Pakistan, Dr. Mansoor Khan Khattak and Dr. Fazl-e-Raziq Durani who were always there as a firm support when ever I needed. I thank to my parents for letting me pursue my dream for so long so far away from home. A special dedication to my grandmother who left this temporary world when I was away from home for studies.
I would like to thank the nana-koza staff and members, and all my Japanese and foreigner friend around for making my stay at Kyoto very much memorable.
Thanks to the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan for supporting my M.Sc and Dr.Eng studies at Kyoto University. Additionally, I highly appreciate the following organizations for their partial support of this project; the grant from ISIJ as an activity of research group, High Precision Process Control via Large Scale Database and Simulation Models, the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (C) 24560940.
Thanks to the Almighty Allah for keeping me healthy and blessing me all the kind and sincere people around.

Publications Extracted from this Dissertation

Journals' Publications:

1. **Iftikhar Ahmad**, Manabu Kano, Shinji Hasebe, Hiroshi Kitada and Noboru Murata: "Prediction of molten steel temperature in steel making process with uncertainty by integrating gray-box model and bootstrap filter", *Journal of Chemical Engineering of Japan*, (submitted)

2. **Iftikhar Ahmad**, Manabu Kano, Shinji Hasebe, Hiroshi Kitada and Noboru Murata: "Gray-Box Modeling for Prediction and Control of Molten Steel Temperature in Tundish", *Journal of Process Control*, (accepted).

3. Toshinori Okura, **Iftikhar Ahmad**, Manabu Kano, Shinji Hasebe, Hiroshi Kitada and Noboru Murata: "High-Performance Prediction of Molten Steel Temperature in Tundish through Gray-Box Model", *ISIJ International*, Vol. 53, No. 1, pp. 76-80, 2013.

Presentation at conferences:

1. **Iftikhar Ahmad**, Manabu Kano, Shinji Hasebe, Hiroshi Kitada and Noboru Murata: "Design of Inner and Outer Gray-Box Models to Predict Molten Steel Temperature in Tundish", *10th International Symposium on Dynamics and Control of Process Systems*, Mumbai, India, December 18-20, 2013.

2. **Iftikhar Ahmad**, Manabu Kano, Shinji Hasebe, Hiroshi Kitada and Noboru Murata: "Comparison of Gray-Box Models to Estimate Molten Steel Temperature", *Program of the 165th ISIJ Meeting*, Tokyo, Japan, March 27-29, 2013. September 10-12, 2012.

3. Shota Sakashita, Toshinori Okura, **Iftikhar Ahmad**, Manabu Kano, Hiroshi Kitada and Noboru Murata: "Gray-Box Model to Control Molten

Steel Temperature in Tundish", *IFAC Workshop on Automation in the Mining, Mineral and Metal Industries*, Gifu, Japan, September 10-12, 2012.